Matriser

Ingmar Björkfelt och Sten Bjon

Åbo 2003

Förord

Linjära samband uttrycks enklast, kortast och tydligast med hjälp av matriser. Att matrismetoder är viktiga beror främst på att ett linjärt samband oftast är den första approximationen av ett mera komplicerat samband. T.ex. kan en krokig kurva $y = f(x)$ ju approximeras **lokalt** av sin tangent, vilken beskrivs av en **linjär** ekvation $\Delta y = f'(x_0)\Delta x$ förutsatt att derivatan $f'(x_0)$ existerar. Då man har en funktion med flera variabler, blir den lokala linjära approximationen mer komplicerad och uttrycks då enklast med hjälp av matriser. I många tillämpningar, t.ex. i ellära, är sambanden färdigt linjära. I dessa fall är den linjära matrisbeskrivningen t.o.m. den fundamentala.

Framställningen av ämnet matriser i detta kompendium strävar till att lära ut både begrepp och metoder. I exemplen visas alltid "små" matriser för att det tydligare skall framgå hur en viss metod fungerar. Man bör emellertid ha klart för sig att metoden i fråga fungerar på samma sätt också för "stora" matriser. Metoderna är oftast algoritmiskt beskrivna så att de utan vidare kan utföras av ett datorprogram. De flesta matematikprogram för datorer innehåller i själva verket färdiga rutiner för större delen av dessa metoder men för att kunna använda rutinerna på ett adekvat sätt bör användaren känna till hur de fungerar. Denna förståelse strävar vi till att ge läsaren genom detta kompendium. För enkelhetens skull behandlas enbart reella matriser även om de flesta begrepp och metoder har sina direkta motsvarigheter för matriser med komplexa element.

Kompendiet är en bearbetning av Ingmar Björkfelts mer än tjugo år gamla kompendium. Alla kapitel har omarbetats. Såväl kapitlens inbördes ordning som ordningsföljden på stoffet inom kapitlen har i någon mån ändrats, nya exempel och övningsuppgifter har tillkommit och några har strukits. Kapitlet om egenvärden och egenvektorer samt kapitlet om diagonalisering är helt nyskrivna. Detsamma gäller avsnittet om spektralframställningen av en symmetrisk matris. De nya avsnitt med finare stil, som finns på några ställen, innehåller stoff som inte nödvändigtvis hör hemma i en grundkurs i ämnet matriser. Bortsett från tilläggen är innehållet i detta förnyade kompendium dock väsentligen detsamma som i det gamla.

STEN BJON

Matematiska institutionen vid Åbo Akademi, Åbo
Den 8 maj 2003

Innehållsförteckning

Tryckfel i kompendiet "Matriser" av Bjon och Björkfelt

Följande fel har hittats i texten (listan är inte uttömmande):

- s. 8: "enligt Sats 2.3 (c)" borde vara "enligt Sats 1.2 (c)"
- s. 74: $\frac{5}{9}x_2^2$ borde vara $\frac{5}{3}x_2^2$.
- s. 114: $\begin{pmatrix} a_1^T \\ \vdots \\ a_p \end{pmatrix}$ borde vara $\begin{pmatrix} a_1^T \\ \vdots \\ a_p^T \end{pmatrix}$ både på raden (5) och (7).
- s. 132: $z = (T \circ S)(x)$ borde vara $z = (S \circ T)(x)$.

Dessutom är följande övningsuppgifter fel formulerade eller har fel i facit: 5.16, 6.19.

1. Gausseliminering

Vi skall till att börja med söka lösningen (lösningarna) till ett så kallat *linjärt ekvationssystem.* Ett sådant system med m ekvationer och n obekanta (m, $n \in \mathbf{Z}_+$) har formen

$$
\begin{array}{c}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
- \quad - \quad - \quad - \quad - \quad - \quad - \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m\,.
\end{array}
\tag{1}
$$

Talen a_{ik} och b_i ($i = 1, \ldots, m$; $k = 1, \ldots, n$ är givna reella tal och symbolerna x_j ($j = 1, \ldots, n$) är så kallade *obekanta* (ibland också kallade *variabler*). Att *lösa* systemet innebär att finna alla n-tiplar $(x_1,\, x_2, \ldots, x_n)$ som satisfierar alla ekvationer i systemet.

Det finns olika metoder för lösning av (1). Den mest praktiska är metoden med Gausseliminering. Det är också den som man använder då man löser ett linjärt ekvationssystem på dator.

Triangulära system och system i echelonform

Låt oss se på ett exempel:

Exempel 1.1. För ekvationssystemet

$$
\begin{aligned}
2x_1 + x_2 - x_3 &= 5 \\
x_2 + x_3 &= 3 \\
3x_3 &= 6
\end{aligned}
$$

är lösningsmetoden självklar: Först löses den tredje ekvationen och man får $x_3 = 2$. Detta utnyttjas sedan i ekvation två, som ger $x_2 = 3 - x_3 = 3 - 2 = 1$. Slutligen ger den första ekvationen, att $x_1 = (5 - x_2 + x_3)/2 = 3$. Man får den entydiga lösningen $(x_1,\, x_2,\, x_3) = (3,\, 1,\, 2)$.

Definition 1.1. Om alla obekanta i ekvationssystem (1) är lika med antalet ekvationer sägs systemet vara *kvadratiskt.* Ett kvadratiskt system sägs vara *uppåt triangulärt* om $a_{ik} = 0$ så snart $k < i$ och *nedåt triangulärt* om $a_{ik} = 0$ så snart $k > i$.

Alla uppåt triangulära system, sådana att $a_{ii} \neq 0$ för alla i, kan som i exempel 1.1 lösas genom s.k. *bakåtsubstitution*: Man börjar med den sista ekvationen, som ger värdet på den sista obekanta, och får tidigare obekanta i tur och ordning genom insättning av redan kända värden. Som de följande exemplen visar behöver systemet ingalunda vara triangulärt för att denna lösningsmetodik skall fungera.

Exempel 1.2. Systemet

$$\begin{aligned} 2x_1 + x_2 - x_3 + x_4 &= 5 \\ x_2 + x_3 - x_4 &= 3 \\ 3x_3 + 6x_4 &= 6 \end{aligned}$$

är inte triangulärt (det är ju inte kvadratiskt). Ger vi däremot x_4 ett godtyckligt värde s ($s \in \mathbf{R}$), som sedan betraktas som bekant, och skriver om systemet,

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 5 - s \\ x_2 + x_3 &= 3 + s \\ 3x_3 &= 6 - 6s\,, \end{aligned}$$

så blir det triangulärt i de obekanta x_1, x_2, x_3 och kan lösas genom bakåtsubstitution. Lösningsmängden blir mängden av alla lösningar

$$\begin{aligned} x_1 &= 3 - 3s \\ x_2 &= 1 + 3s \quad (s \in \mathbf{R}) \\ x_3 &= 2 - 2s\,. \end{aligned}$$

Vi säger att x_4 är en *fri variabel*, eftersom vi ju valde denna godtyckligt (fritt). De övriga variablerna x_1, x_2 och x_3 är *basvariabler*.

Exempel 1.3. Systemet

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 - x_5 &= 2 \\ x_3 + x_4 + x_5 &= -1 \\ x_4 - 2x_5 &= 4 \end{aligned}$$

är triangulärt i variablerna x_1, x_3 och x_4. De två övriga får bli fria variabler och tilldelas godtyckliga värden: $x_2 = s$, $x_5 = t$ ($s, t \in \mathbf{R}$). Lösningen, som fås genom bakåtsubstitution, blir

$$\begin{array}{rrrrl} x_1 = & 13 & -2s & +8t & \\ x_2 = & -5 & & -3t\,, & \quad (s, t \in \mathbf{R}) \\ x_3 = & 4 & & +2t & \end{array}$$

Systemen i exemplen 1.1-1.3 sägs ha s.k. echelonform:

Definition 1.2. Ett system har *echelonform* om den första icke-noll-termen (dvs. en term med koefficienten $a_{ij} \neq 0$) i varje ekvation ligger längre till höger än i föregående ekvation. Den första koefficienten olik noll i varje rad sägs då vara ett *pivotelement*.

Exempel 1.4. Varje uppåt triangulärt system är uppenbarligen i echelonform.

Det är klart att **högst ett pivotelement är associerat med varje obekant** och att det **finns högst lika många pivotelement som det finns rader**. De obekanta (eller variabler) som är förbundna med något pivotelement är just de som vi i exemplen kallat *basvariabler* medan de övriga kallats *fria variabler*. Varje system i echelonform kan skrivas som ett triangulärt system i sina basvariabler genom att man flyttar alla termer som innehåller fria variabler till ekvationernas högra led.

Omformning till echelonform

Definition 1.3. Två ekvationssystem sägs vara *ekvivalenta* om de har samma lösningsmängd.

Anmärkning. Denna definition på ekvivalens är den mest praktiska för våra behov. Man kan visa att två system (S_1) och (S_2) är ekvivalenta i ovanstående mening om och endast om de är ekvivalenta i logisk mening: $(S_1) \iff (S_2)$. Beviset för detta kräver emellertid en väsentlig del av den teori som vi står i beråd att utveckla.

Vår strävan är nu att skriva ett linjärt ekvationssystem i form av ett echelonsystem, som är ekvivalent med det givna systemet. Lösningarna kan ju sedan fås genom bakåtsubstitution. Metodiken framgår genom några exempel.

Exempel 1.5. Betrakta systemet

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \quad (1) \\ 2x_1 + 5x_2 + x_3 &= 9 \quad (2) \\ x_1 + 4x_2 - 6x_3 &= 1 \quad (3)\,, \end{aligned}$$

där vi har numrerat ekvationerna för att lättare kunna förklara vad vi gör. Först bildar vi ett nytt system genom att *eliminera* x_1 ur (2) och (3) med hjälp av (1). Detta tillgår så, att vi multiplicerar bägge leden i (1) med talet 2 och subtraherar resultatet från (2). Då försvinner den term som innehåller x_1 ur (2). På samma sätt subtraherar vi 1 gånger (1) från (3), varvid termen innehållande x_1 försvinner ur (3). Dessa operationer kan i symbolspråk uttryckas genom

$$\begin{aligned} (2) &\to (2) - 2(1) \\ (3) &\to (3) - 1(1)\,, \end{aligned}$$

där t.ex. den första raden kan utläsas "ersätt rad (2) med rad (2) minus 2 gånger rad (1)". Det nya systemet blir nu

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \quad (1') \\ x_2 - 5x_3 &= -3 \quad (2') \\ 2x_2 - 9x_3 &= -5 \quad (3')\,. \end{aligned}$$

Det system som bildas av $(2')$ och $(3')$ behandlas sedan enligt samma mönster, dvs. x_2 elimineras ur $(3')$ genom att man utför operationen

$$(3') \to (3') - 2(2')\,,$$

och man får ett system i echelonform (det är t.o.m. triangulärt):

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \\ x_2 - 5x_3 &= -3 \\ x_3 &= 1\,. \end{aligned}$$

Lösningen fås nu genom bakåtsubstitution och visar sig bli $(x_1, x_2, x_3) = (-1, 2, 1)$.

Den operation som vi använt i föregående exempel kallar vi

Basoperation 1 ($BO1$): *Subtrahera en multipel av en ekvation från en* annan *ekvation:*

$$(i) \to (i) - \lambda(k) \quad (i \neq k)\,.$$

Denna basoperation har vi anledning att dela upp i två litet strängare delar:

($BO1^+$) *Subtrahera en multipel av en ekvation från en* senare *ekvation.*

($BO1^-$) *Subtrahera en multipel av en ekvation från en* tidigare *ekvation.*

Andra operationer, som uppenbarligen kan användas på ett system utan att systemets lösningmängd ändras, är

Basoperation 2 ($BO2$): *Två ekvationer, säg* (i) *och* (k)*, byter plats:*

$$(i) \leftrightarrow (k)\,.$$

Basoperation 3 ($BO3$): *En ekvation, säg ekvation* (i)*, multipliceras med ett tal* λ *som inte är* 0*:*

$$(i) \to \lambda(i)\,.$$

Anmärkning. Basoperationerna $(BO1)-(BO3)$ är beroende av varandra. Man finner t.ex. lätt att ($BO2$) kan härledas ur ($BO1$) och ($BO3$), vilket ju gör ($BO2$) överflödig ur teoretisk synvinkel. Vid praktiska kalkyler är det emellertid bäst att ha tillgång till dessa tre enkla regler.

En viktig sak återstår att verifiera, nämligen

Sats 1.1. *De tre basoperationerna* $(BO1)-(BO3)$*, tillämpade på ett ekvationssystem, ger upphov till ett system som är ekvivalent med det ursprungliga.*

Bevis. Om man låter två ekvationer byta plats i ett system, så går det ju att byta tillbaka, och om man multiplicerar en ekvation med λ ($\neq 0$), så kan man sedan multiplicera samma ekvation med $1/\lambda$ och på det sätter återställa det ursprungliga systemet. Det är därför trivialt att användning av ($BO2$) och ($BO3$) ger upphov till nya system, som är ekvivalenta med det ursprungliga. Nästan lika enkelt är det att bevisa satsens påstående för ($BO1$):

Om vi tillämpar operationen $(k) \to (k) - \lambda(i)$ på systemet

$$(S_1) \qquad \begin{cases} \quad \vdots \\ V_i = b_i \quad (i) \\ \quad \vdots \\ V_k = b_k \quad (k) \\ \quad \vdots \end{cases} ,$$

där de vänstra leden kort betecknats med V_j, så får vi

$$(S_2) \qquad \begin{cases} \qquad \vdots & \\ \qquad V_i = b_i & (i') \\ \qquad \vdots & \\ V_k - \lambda V_i = b_k - \lambda b_i & (k') \\ \qquad \vdots & . \end{cases}$$

Och omvänt, om vi använder operationen $(k') \to (k') + \lambda(i')$ på (S_2), så återfår vi (S_1). Således är systemen (S_1) och (S_2) ekvivalenta. $\diamondsuit$

Vi kan sammanfatta proceduren för hur man överför ett system på echelonform i en algoritm:

Algoritm för echelonform:

1. Permutera ekvationerna (dvs. använd $(BO2)$), så att koefficienten för den 1:a variabeln i den 1:a ekvationen är olik 0, om detta är möjligt. I annat fall, betrakta nästa variabel som den första och gör samma sak om det är möjligt, osv.
2. Eliminera den första variabeln ur de övriga ekvationerna (använd $(BO1^+)$).

3*a*. Upprepa genom att tillämpa 1. och 2. på det mindre system som uppstår då man bortser från den första ekvationen.

3*b*. Sluta då det mindre systemet innehåller bara en ekvation eller ingen variabel alls.

Räkneschemat

Då man Gausseliminerar är det faktiskt bara koefficienterna framför variablerna och talen i högerleden som spelar någon roll. Man kan i själva verket utföra elimineringen helt och hållet i ett räkneschema (= matris), som innehåller enbart dessa tal:

Exempel 1.6. Om systemet

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 2 \\ 2x_1 + 6x_2 + 9x_3 + 5x_4 &= 4 \\ -x_1 - 3x_2 + 3x_3 \qquad\quad &= 1 \end{aligned}$$

skall omformas till echelonform, skriver man det först som ett räkneschema

$$(2) \qquad S = \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 2 \\ 2 & 6 & 9 & 5 & 4 \\ -1 & -3 & 3 & 0 & 1 \end{array}\right).$$

Utgående från (2) ger nu elimineringen följande räknescheman, där en pil antyder att basoperationer har använts för att få fram nästa räkneschema:

$$S \overset{BO1^+}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 2 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 6 & 2 & 3 \end{array}\right) \overset{BO1^+}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 3 & 2 & 2 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right).$$

Fullständigt utskriven blir echelonformen alltså

$$\begin{aligned} x_1 + 3x_2 + 3x_3 + 2x_4 &= 2 \\ 3x_3 + x_4 &= 0 \\ 0 &= 3\,. \end{aligned}$$

Den tredje ekvationen är inkonsistent, dvs. falsk (se definitionen nedan). Systemet saknar alltså lösning.

Definition 1.4. En ekvation eller ett ekvationssystem är *konsistent* om (minst) en lösning existerar och *inkonsistent* om lösningar saknas.

Exempel 1.7. Vi skall lösa systemet

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 - x_2 + x_3 &= 2 \\ x_2 + x_3 &= 1 \\ x_1 - x_2 + x_3 &= 2 \end{aligned}$$

och överför det först i echelonform,

$$\begin{aligned} &\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & -1 & 1 & 2 \end{array}\right) \overset{BO1^+}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -2 & 0 & 1 \end{array}\right) \\ (3)\qquad &\overset{BO1^+}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 \\ 0 & 0 & 2/3 & 1 \\ 0 & 0 & 2/3 & 1 \end{array}\right) \overset{BO1^+}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 \\ 0 & 0 & 2/3 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) = S\,. \end{aligned}$$

Det sista räkneschemet S svarar mot echelonformen

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ -3x_2 - x_3 &= 0 \\ \frac{2}{3}x_3 &= 1 \\ 0 &= 0\,. \end{aligned}$$

Lösningen fås sedan genom bakåtsubstitution, dvs. vi sätter in värdet $x_3 = 3/2$ i den andra ekvationen, vilket ger x_2 osv. Men **precis samma bakåtsubstitutionen kan också göras i ett räkneschema** med hjälp av $(BO1^-)$. Man utgår då från det sista räkneschemat S i (3) och börjar med att eliminera koefficienterna för x_3 ur de två första ekvationerna med hjälp av den tredje:

$$\begin{aligned} &S \overset{BO3}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 \\ 0 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 0 \end{array}\right) \\ &\overset{BO1^-}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 0 & -1/2 \\ 0 & -3 & 0 & 3/2 \\ 0 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 0 \end{array}\right) \overset{BO1^-}{\rightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 0 \end{array}\right). \end{aligned}$$

Lösningen kan avläsas längst till höger i det sista schemat: $x_1 = 0$, $x_2 = -1/2$ och $x_3 = 3/2$.

Det sista räkneschemat är av en alldeles speciell form, reducerad echelonform:

Definition 1.4. Ett räkneschema (= matris) som svarar mot ett system i echelonform kallas en *echelonmatris.* En echelonmatris, i vilken alla pivotelement är ettor och dessa ettor är de enda som inte är noll i sin kolonn (= kolumn) sägs vara i *reducerad echelonform.*

Exempel 1.8. T.ex. det sista räkneschemat i exempel 1.7 och t.ex. varje räkneschema av formen

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

där varje stjärna står för ett godtyckligt tal, är i reducerad echelonform. De variabler som svarar mot pivotkolonner, dvs. x_2, x_4 och x_5, är basvariabler medan de övriga, x_1, x_3 och x_6 är fria variabler.

Lösning av ekvationssystem

Väl motiverade av de tre senaste exemplen kan vi skriva ut en algoritm:

Algoritm för lösning av ett ekvationssystem:

1. Omforma systemet till echelonform (enligt den algoritm som vi har beskrivit tidigare).
2. Sluta om någon ekvation är inkonsistent (systemet saknar lösningar). I annat fall, fortsätt med bakåtsubstitution dvs. med omformning av räkneschemat till reducerad echelonform med hjälp av $(BO1^-)$ och $(BO3)$.
   a. Om det inte finns fria variabler fås en entydig lösning.
   b. Om det finns fria variabler, så kan basvariablerna skrivas som funktioner av de fria varablerna (de senare ges godtyckliga värden $s, t, \ldots$ och flyttas över till det högra ledet). Det finns alltså oändligt många lösningar.

I fortsättningen använder vi beteckningarna #(ekvationer), #(fria variabler), etc. för antalet ekvationer, antalet fria variabler, etc.

Vi sammanfattar:

Sats 1.2. *Följande gäller för ett linjärt ekvationssystem:*

$$(a) \quad \left\{\begin{array}{c} \textit{Systemet konsistent} \\ \#\textit{(fria variabler)} = 0 \end{array}\right\} \Longrightarrow \textit{Lösningen är entydig};$$

$$(b) \quad \left\{\begin{array}{c} \textit{Systemet konsistent} \\ \#\textit{(fria variabler)} > 0 \end{array}\right\} \Longrightarrow \#\textit{(lösningar)} = \infty\,;$$

$$(c) \quad \#\textit{(fria variabler)} + \#\textit{(basvariabler)} = \#\textit{(variabler)}\;;$$

$$(d) \quad \#\textit{(basvariabler)} = \#\textit{(pivotelement)};$$

$$(e) \quad \#\textit{(basvariabler)} \leq \#\textit{(ekvationer)}.$$

Definition 1.5. Ett linjärt ekvationssysten är *homogent* om varje högerled är en nolla.

Sats 1.3. *Ett homogent system*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ - \quad - \quad - \quad - \quad - \quad - \quad & - \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

är alltid konsistent, ty det har åtminstone lösningen $x_1 = x_2 = \cdots = x_n = 0$ (den s.k. triviala lösningen*). Om $m < n$ har systemet oändligt många lösningar.*

Bevis. Det är uppenbart att den triviala lösningen alltid är en lösning till ett homogent system. Antag att $m < n$. Då är $\#$(basvariabler) $\leq m < n$ och enligt Sats 2.3 (c) $\#$(fria variabler) > 0. Eftersom det finns fria variabler, är antalet lösningar oändligt. $\diamondsuit$

Då man löser ett homogent ekvationssystem är det onödigt att skriva ut nollorna i högerledet i räkneschemat. Dessa nollor förblir ju nollor då man använder basoperationer.

Exempel 1.9. Vid lösning av det homogena systemet

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ 3x_1 + 6x_2 + 10x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

gör man alltså följande kalkyler med räknescheman

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 10 \\ 1 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Eftersom variabeln x_2 är fri, sätter vi $x_2 = s$, där s representerar ett godtyckligt tal, och finner lösningarna $x_1 = -2s$, $x_2 = s$, $x_3 = 0$ $(s \in \mathbf{R})$.

Övningsuppgifter

1. Lös systemet

$$\begin{aligned} u + v + w &= -2 \\ 3u + 3v - w &= 6 \\ u - v + w &= -1\,. \end{aligned}$$

2. Lös systemet

$$\begin{aligned} 2x_1 + x_2 + x_3 + x_4 &= 2 \\ x_2 + x_3 + x_6 &= 3 \\ x_3 &= -1 \\ x_5 + x_6 &= 4\,. \end{aligned}$$

3. Skriv följande system i echelonform

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 + 3x_3 + 4x_4 &= 2 \\ x_1 + 2x_2 + 4x_3 + 5x_4 &= 3 \\ x_1 + 2x_2 + 4x_3 + 6x_4 &= 4\,. \end{aligned}$$

4. Lös med Gausseliminering

$$\begin{aligned} x_1 + x_2 - x_3 + x_4 &= 17 \\ x_1 + 2x_2 + 2x_4 &= 15 \\ x_1 + x_2 &= 17\,. \end{aligned}$$

5. Lös

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + 2x_2 + x_3 &= 3 \\ x_2 + 2x_3 + x_4 &= 3 \\ x_3 + 2x_4 + x_5 &= 1 \\ x_4 + 2x_5 + x_6 &= -3 \\ x_5 + 2x_6 &= -5\,. \end{aligned}$$

6. Två sjöar A och B är förbundna med en kanal. Under en viss tidsperiod (ett år) simmar 10% av fiskarna från A till B medan 20% av fiskarna i B simmar till A. Om det vid årets slut fanns 20 000 fiskar i A och 3 000 i B, hur många fanns det vid årets början i A respektive B? (Vi antar att antalet fiskar, som föds och dör i respektive sjö, är lika.)
7. Om fiskarna i A och B har samma "rörlighet" som i uppgift 6 och det dessutom är så att antalet fiskar i respektive sjö är lika vid årets början och vid årets slut, vad är förhållandet mellan antalet fiskar i A och antalet fiskar i B?
8. Tre bakteriearter lever samtidigt i ett provrör och livnär sig på tre slag av födoämnen. Antag att en bakterie av en art i ($i = 1, 2, 3$) konsumerar i medeltal a_{ki} enheter av födoämnet k ($k = 1, 2, 3$) per dag. Antag att $a_{11} = 1$, $a_{12} = a_{21} = 1$, $a_{13} = a_{31} = 1$, $a_{22} = 2$, $a_{23} = a_{32} = 3$ och $a_{33} = 5$. Antag vidare att det tillförs 15 000 enheter av det första födoämnet, 30 000 enheter av det andra och 45 000 av det tredje per dag. Hur många individer av de tre arterna kan leva i provröret, då vi antar att all föda konsumeras?
9. Undersök hur många lösningar ekvationssystemet

$$\begin{aligned} x + ay &= a - 2 \\ (2-a)x - 3y &= -1 \end{aligned}$$

har för olika värden på den reella konstanten a. Ange lösningarna.

10. För vilka värden på a har systemet

$$\begin{aligned} (a+2)x + (a+1)y + az &= 0 \\ x - z &= 0 \\ ax + (2a-1)y + (a-2)z &= 0 \end{aligned}$$

icke-triviala lösningar?

11. Antag att de tre punkterna $(1, -5)$, $(-1, 1)$ och $(2, 7)$ ligger på en parabel $y = ax^2 + bx + c$. Bestäm a, b och c genom att lösa ett linjärt ekvationssystem.
12. Ett arv på 24 000 euro fördelades på tre fonder så, att den andra fonden får dubbelt så mycket som första. Fonderna ger en årlig ränta på 9%, 10% respektive 6% och pengarna i alla tre fonderna avkastar tillsammans 2 210 euro under det första året. Hur mycket pengar sattes i varje fond?
13. Visa att $(BO2)$ följer ur $(BO1)$ och $(BO3)$.

2. Matriser och vektorer

Definition 2.1. En *matris* är ett rektangulärt schema

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

av reella tal, som kallas *matriselement*. Matrisen A har m *rader* och n *kolonner* och sägs därför vara en m/n*-matris*, där m/n utgör matrisens *typ*.

Då man vill framhålla beteckningen för A:s matriselement, skriver man ofta $A = (a_{ik})$. Talet a_{ik} är alltså matriselementet i rad i och kolonn k eller som man säger matriselementet "på platsen (i, k)".

Definition 2.2. Två matriser $A = (a_{ik})$ och $B = (b_{ik})$ sägs vara *lika* ($A = B$), om de är av samma typ och $a_{ik} = b_{ik}$ för varje i och k.

Exempel 2.1. Ett räkneschema som t.ex. det här till vänster,

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 10 \\ 4 & 5 & 6 & 11 \\ 7 & 8 & 9 & 12 \end{array}\right), \qquad \begin{pmatrix} 0 & -2 \\ 5 & 3 \\ -1 & -3 \end{pmatrix},$$

är en 3/4-matris. Att vi har skrivit in streck som utmärker var likhetstecknen skall stå, hindrar inte att vi kallar schemat en matris. Till höger har vi en 3/2-matris.

Definition 2.3. För varje typ m/n kan vi skriva ut en *nollmatris*, dvs. en matris som består av enbart nollor. Dessa betecknas alla med 0. En *radvektor* respektive *kolonnvektor* är en matris av formen

$$\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}, \qquad \text{respektive} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

En m/n-matris sägs vara *kvadratisk* om $m = n$. En kvadratisk matris $A = (a_{ik})$ är en *diagonalmatris* om $a_{ik} = 0$ för alla i och k med $i \neq k$, dvs. om den har utseendet

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ - & - & - & - \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Elementen a_{ii}, $i = 1, \ldots, n$, i en kvadratisk matris bildar matrisens *(huvud)diagonal*. En *enhetsmatris* är en diagonalmatris med enbart ettor i diagonalen.

Alla enhetsmatriser betecknas med I. Då man vill framhålla enhetsmatrisens typ n/n, kan man skriva I_n. *Kroneckers delta*,

$$\delta_{ik} = \begin{cases} 1, & \text{om } i = k \\ 0, & \text{om } i \neq k \end{cases}$$

är den gängse beteckningen för matriselementen i en enhetsmatris: $I = (\delta_{ik})$. En *uppåt* respektive *nedåt triangulär* matris har den allmänna formen

$$\begin{pmatrix} * & * & \cdots & * \\ & * & \cdots & * \\ - & - & - & - \\ & & & * \end{pmatrix} \quad \text{resp.} \quad \begin{pmatrix} * & & & \\ * & * & & \\ - & - & - & - \\ * & * & \cdots & * \end{pmatrix},$$

där vi har använt konventionen att **en stjärna symboliserar vilket tal som helst** medan **ett tomrum står för en nolla**.

Addition och multiplikation med skalär

Vi skall nu definiera två räkneoperationer för matriser. Antag att $A = (a_{ik})$ och $B = (b_{ik})$ är två matriser av samma typ m/n och låt λ vara ett reellt tal, en *skalär*.

Definition 2.2. *Addition* av matriser definieras genom

$$A + B = (a_{ik} + b_{ik}),$$

dvs. matriser adderas så att motsvarande matriselement adderas. *Multiplikation med skalär* definieras genom

$$\lambda A = (\lambda a_{ik}),$$

dvs. en matris multipliceras med en skalär λ så att varje matriselement multipliceras med λ.

Exempel 2.2. Vi har t.ex. att

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix},$$

$$(-2)\begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ -6 & 8 \end{pmatrix}.$$

Följande egenskaper (räkneregler) kan lätt verifieras för addition och multiplikation med skalär: Om A, B, C är m/n-matriser och α och β är reella tal, dvs. skalärer, så gäller:

(i) $A + B = B + A$ (kommutationslagen);
$(A + B) + C = A + (B + C)$ (associationslagen);

$A + 0 = A$;

Mot varje matris A svarar en matris A', sådan att $A + A' = 0$ (välj $A' = (-1)A$).

(*ii*) $(\alpha\beta)A = \alpha(\beta A)$;

$1\,A = A$;

$0\,A = 0$;

$\alpha 0 = 0$.

(*iii*) $\alpha(A + B) = \alpha A + \alpha B$ (distributionslag);

$(\alpha + \beta)A = \alpha A + \beta A$ (distributionslag).

Som ett exempel verifierar vi kommutationslagen för matriser med hjälp av kommutationslagen för reella tal: Om $A = (a_{ik})$ och $B = (b_{ik})$, är uppenbarligen

$$A + B = (a_{ik} + b_{ik}) = (b_{ik} + a_{ik}) = B + A\,.$$

Matrismultiplikation

Vi inför ännu en räkneoperation för matriser, nämligen multiplikation av matriser med varann. Ett exempel motiverar de definitioner som skall slås fast:

Exempel 2.3. Betrakta ett linjärt ekvationssystem

$$(1) \qquad \begin{aligned} 3x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 - x_2 + x_3 &= 4 \\ x_1 + x_2 - x_3 &= 6\,. \end{aligned}$$

Med detta system är det naturligt att associera matrisen

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

den s.k. *koefficientmatrisen*, och kolonnvektorerna

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 5 \\ 4 \\ 6 \end{pmatrix}.$$

Vi vill definiera matrisprodukten $A\mathbf{x}$ av A och $\mathbf{x}$ så, att systemet (1) kan skrivas $A\mathbf{x} = \mathbf{b}$. Detta betyder att vår definition måste bli sådan att

$$A\mathbf{x} = \begin{pmatrix} 3 & 2 & 3 \\ 2 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 + 3x_3 \\ 2x_1 - 1x_2 + 1x_3 \\ 1x_1 + 1x_2 - 1x_3 \end{pmatrix}.$$

För att få t.ex. det andra elementet $2x_1 - x_2 + x_3$ i matrisprodukten $A\mathbf{x}$, väljer vi den andra raden $\bar{\mathbf{a}}_2$ i A och bildar summan av produkterna av element på motsvarande platser i $\bar{\mathbf{a}}_2$ och $\mathbf{x}$.

Exemplet visar hur definitionerna bör göras:

Definition 2.3. Produkten av en radvektor $\mathbf{a}$ och en kolonnvektor $\mathbf{b}$ med lika många element är en 1/1-matris definierad genom

$$\mathbf{a}\,\mathbf{b} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = (a_1 b_1 + a_2 b_2 + \ldots + a_n b_n)\,.$$

Man identifierar sedan 1/1-matrisen $(a_1 b_1 + \cdots + a_n b_n)$ med talet $a_1 b_1 + \cdots + a_n b_n$.

Definition 2.4. Produkten $A\mathbf{x}$ av en

$$m/n\text{-matris } A = (a_{ik}) \text{ och en kolonnvektor } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ av typ } n/1,$$

är en kolonnvektor av typ $m/1$, i vilken element nummer i fås som produkten $\bar{\mathbf{a}}_i\mathbf{x}$ av rad i i A med $\mathbf{x}$:

$$A\mathbf{x} = \begin{pmatrix} \bar{\mathbf{a}}_1\mathbf{x} \\ \vdots \\ \bar{\mathbf{a}}_m\mathbf{x} \end{pmatrix}.$$

Om vi betecknar element nummer i i $A\mathbf{x}$ med $(A\mathbf{x})_i$, är alltså

$$(A\mathbf{x})_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \sum_{k=1}^{n} a_{ik}x_k \quad (i = 1, \ldots, m)\,.$$

Det är här på sin plats att göra en nyttig observation. I exempel 2.4 ovan får vi enligt reglerna för addition och multiplikation med skalär

$$\begin{aligned} A\mathbf{x} &= \begin{pmatrix} 3x_1 + 2x_2 + 3x_3 \\ 2x_1 - 1x_2 + 1x_3 \\ 1x_1 + 1x_2 - 1x_3 \end{pmatrix} = \begin{pmatrix} 3x_3 \\ 2x_1 \\ 1x_1 \end{pmatrix} + \begin{pmatrix} 2x_2 \\ -1x_2 \\ 1x_2 \end{pmatrix} + \begin{pmatrix} 3x_3 \\ 1x_3 \\ -1x_3 \end{pmatrix} \\ &= x_1 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}. \end{aligned}$$

$A\mathbf{x}$ är en så kallad *linjärkombination* av kolonnerna i A. Koefficienten framför varje kolonn är motsvarande komponent i vektorn $\mathbf{x}$.

Allmänt bevisas på samma sätt som ovan:

Sats 2.1. *Om* $A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$ *är en m/n-matris med kolonnerna* $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$ *och* $\mathbf{x}$ *är en kolonnvektor med komponenterna* x_1, x_2, …, x_n, *så gäller att* $A\mathbf{x}$ *är en* linjärkombination,

$$A\mathbf{x} = \mathbf{x}_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n\,,$$

av $\mathbf{a}_1$, $\mathbf{a}_2$,…, $\mathbf{a}_n$.

Låt oss nu anta att A är en m/n-matris och att $B = (\,\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_p\,)$ är en n/p-matris. Då är B:s kolonner $\mathbf{b}_j$ av typen $n/1$, så att produkterna $A\mathbf{b}_j$ kan bildas.

Definition 2.5. Med matrisprodukten AB av de ovannämnda matriserna avses den matris vars kolonner (i ordningsföljd) är $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$, dvs.

$$AB = (\,A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p\,)\ .$$

Hur ser matriselementet $(AB)_{ik}$ på platsen (i, k) i matrisen AB ut? Enligt definitionen är $(AB)_{ik}$ element nummer i i kolonnen $A\mathbf{b}_k$, dvs. produkten av **rad** i i A med **kolonn** k i B $(= \mathbf{b}_k)$,

$$\begin{aligned}(AB)_{ik} &= (\,a_{i1} \quad \cdots \quad a_{in}\,)\begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk}\end{pmatrix} \\ &= a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}\ .\end{aligned}$$

Vi sammanfattar tre sätt att se på matrismultiplikationen i en sats:

Sats 2.2. *Låt A vara en m/n-matris och B en n/p-matris. Vi betecknar raderna i A med $\bar{\mathbf{a}}_i$ och kolonnerna i B med $\mathbf{b}_k$ så att*

$$A = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_m \end{pmatrix} \qquad \textit{och} \qquad B = (\,\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_p\,)\ .$$

Då gäller:

$$\begin{aligned} &(i) \qquad (AB)_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}\,; \\ &(ii) \qquad AB = (\,A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p\,)\,; \\ &(iii) \qquad AB = \begin{pmatrix} \bar{\mathbf{a}}_1 B \\ \vdots \\ \bar{\mathbf{a}}_m B \end{pmatrix}. \end{aligned}$$

Bevis. Vi har redan verifierat (i) medan (ii) är själva definitionsuttrycket. Då återstår det bara att verifiera (iii): Enligt definitionerna 2.5 och 2.4 är

$$AB = (\,A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p\,) = \begin{pmatrix} \bar{\mathbf{a}}_1\mathbf{b}_1 & \cdots & \bar{\mathbf{a}}_1\mathbf{b}_p \\ \vdots & & \vdots \\ \bar{\mathbf{a}}_m\mathbf{b}_1 & \cdots & \bar{\mathbf{a}}_m\mathbf{b}_p \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{a}}_1 B \\ \vdots \\ \bar{\mathbf{a}}_m B \end{pmatrix}.\ \Diamond$$

Anmärkning. Observera att produkten av två matriser A och B av typen m/n respektive p/q är definierad bara om $n = p$ och att AB då detta gäller kommer att vara av typen m/q. Om $n \neq p$ så är produkten AB alltså inte definierad.

Exempel 2.4. En produkt av två matriser av typerna 3/2 och 2/2 är definierad och resultatet blir en 3/2-matris:

$$\begin{pmatrix} 2 & 1 \\ 4 & 3 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 2\cdot 5 + 1\cdot 7 & 2\cdot 6 + 1\cdot 8 \\ 4\cdot 5 + 3\cdot 7 & 4\cdot 6 + 3\cdot 8 \\ -2\cdot 5 - 1\cdot 7 & -2\cdot 6 - 1\cdot 8 \end{pmatrix} = \begin{pmatrix} 17 & 20 \\ 41 & 48 \\ -17 & -20 \end{pmatrix}.$$

Vi kommer i fortsättningen att behöva en formel rörande summatecken. Giltigheten hos denna formel beror i grunden på att termernas ordningföljd i en summa av tal kan bytas om hur som helst och på att termer kan grupperas med parenteser hur som helst:

Betrakta mn stycken dubbelindicerade tal, som vi skriver i form av ett rektangulärt schema:

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ - & - & - & - \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix}.$$

Om vi först summerar talen i kolonnerna och därefter bildar summan av kolonnsummorna, så får vi

$$\sum_{i,k} a_{ik} = \sum_{k=1}^{n} \left(\sum_{i=1}^{m} a_{ik} \right).$$

Å andra sidan, om vi först summerar talen i raderna och därefter bildar summan av radsummorna, får vi

$$\sum_{i,k} a_{ik} = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} \right).$$

Eftersom dessa två summor är lika, fås formeln

$$\sum_{k=1}^{n} \left(\sum_{i=1}^{m} a_{ik} \right) = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} \right),$$

som löst uttryckt säger att ordningsföljden på summatecken kan bytas om.

Beteckningen M_{ik} för matriselementet på platsen (i,k) i en matris M är ofta bekväm att använda. Vi använder den i beviset av nästa sats.

Sats 2.3. (*a*) *Om matrisen A är av typen m/n, B av typen n/p och C av typen p/r, så gäller* associationslagen

$$A(BC) = (AB)C.$$

(*b*) *Om matrisen A är av typen m/n samt B och C av typen n/p, så gäller* distributionslagen

$$A(B+C) = AB + AC.$$

Bevis. (*a*) För ett matriselement på en godtycklig plats (i,k) i $A(BC)$ är

$$\begin{aligned}
(A(BC))_{ik} &= \sum_{j=1}^{n} A_{ij}(BC)_{jk} = \sum_{j=1}^{n} A_{ij}\left(\sum_{l=1}^{p} B_{jl}C_{lk}\right) \\
&= \sum_{j=1}^{n}\sum_{l=1}^{p} A_{ij}B_{jl}C_{lk} = \sum_{l=1}^{p}\sum_{j=1}^{n} A_{ij}B_{jl}C_{lk} \\
&= \sum_{l=1}^{p}\left(\sum_{j=1}^{n} A_{ij}B_{jl}\right) C_{lk} = \sum_{l=1}^{p}(AB)_{il}C_{lk} = ((AB)C)_{ik} \ ,
\end{aligned}$$

dvs. $A(BC) = (AB)C$.

(*b*) För ett matriselement på en godtycklig plats (i,k) i $A(B+C)$ är

$$\begin{aligned}
(A(B+C))_{ik} &= \sum_{j=1}^{n} A_{ij}(B+C)_{jk} = \sum_{j=1}^{n} A_{ij}(B_{jk}+C_{jk}) \\
&= \sum_{j=1}^{n}(A_{ij}B_{jk}+A_{ij}C_{jk}) = \sum_{j=1}^{n} A_{ij}B_{jk} + \sum_{j=1}^{n} A_{ij}C_{jk} = (AB+AC)_{ik} \ ,
\end{aligned}$$

dvs. $A(B+C) = AB + AC$. $\diamondsuit$

Märk 1. Multiplikation med en enhetsmatris är speciellt enkel, dvs.

$$IA = A \quad \text{och} \quad AI = A$$

om matrisprodukterna är definierade.

Märk 2. Matrismultiplikationen behöver **inte** vara **kommutativ**. Om A är av typen m/n och B av typen n/p, så är AB definierat medan BA inte är definierat om $m \neq p$. Men också då $m = p$ kan AB och BA vara olika, vilket följande exempel visar:

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

Märk 3. En matrisprodukt kan kan bli noll utan att någondera faktorn är noll:

$$AB = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 6 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

På grund av detta behöver förkortningsregeln inte gälla, dvs. ur $AB = AC$ behöver inte följa att $B = C$. Med samma matriser A och B som nyss kan vi välja $C = 0$. Då är ju $AB = 0 = AC$ men $B \neq C$.

Transponerade matriser

Låt A vara en godtycklig m/n-matris.

Definition 2.6. Med A:s *transponerade* matris A^T (ofta uttalat "A-transponat") förstås den matris som fås då man i A låter rader och kolonner byta plats.

Om A är av typen m/n, är A^T av typen n/m. Vidare gäller för matriselementen i dessa matriser att $(A^T)_{ik} = A_{ki}$.

Exempel 2.5. Om

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad \text{så är} \quad A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

För transponeringsoperationen gäller följande räkneregler:

Sats 2.4. *Om A och B är matriser av sådan typ att ifrågakommande operationer är definierade och om $\lambda \in \mathbf{R}$, är*

$$\begin{aligned}
&(i) && (A+B)^T = A^T + B^T, \\
&(ii) && (\lambda A)^T = \lambda A^T, \\
&(iii) && (AB)^T = B^T A^T, \\
&(iv) && (A^T)^T = A.
\end{aligned}$$

Bevis. Vi bevisar t.ex. regel (iii): Matriselementet på en godtycklig plats (k, i) i $(AB)^T$ är

$$\begin{aligned}
((AB)^T)_{ki} = (AB)_{ik} &= \sum_j A_{ij} B_{jk} = \sum_j (A^T)_{ji} (B^T)_{kj} \\
&= \sum_j (B^T)_{kj} (A^T)_{ji} = (B^T A^T)_{ki}.
\end{aligned}$$

Således gäller regeln (iii). $\diamondsuit$

Anmärkning. En följd av detta är en annan variant av distributionslagen än den i Sats 2.3: Om matriserna A, B och C är sådana att ifrågakommande operationer är definierade, är

$$(A+B)C = AC + BC. \tag{2}$$

Enligt Sats 2.3 är nämligen $C^T(A^T + B^T) = C^T A^T + C^T B^T$ och genom att transponera bägge leden och tillämpa reglerna i Sats 2.4 fås (2).

Övningsuppgifter

1. Lös matrisekvationen

$$\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Räkna också ut $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}$.

2. Bestäm alla kolonnvektorer $\mathbf{x}$ sådana att $A\mathbf{x} = \mathbf{b}$, där

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

3. Lös matrisekvationen $AX = B$, där

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 5 \end{pmatrix} \quad \text{och} \quad B = \begin{pmatrix} 1 & 2 \\ 5 & 3 \end{pmatrix}.$$

4. Låt x_n och y_n beteckna invånarantalet i centrum respektive förorterna i en stad år n. Årligen flyttar 3% av mänskorna i centrum ut till förorterna och 2% av förortsbefolkningen in till centrum. Dessutom flyttar 1% av hela invånartalet från orten (från varje stadsdel) varje år medan 500 flyttar till staden. Av de sistnämnda bosätter sig 50 i centrum medan 450 bosätter sig i förorterna. Bestäm de rekursionsformler, som ger x_{n+1} och y_{n+1} uttryckta med hjälp av x_n och y_n samt skriv dessa i matrisform (A är en 2/2-matris):

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = A\begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

5. Bestäm mängden $\mathcal{M}$ av alla matriser X som kommuterar med $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, dvs. som har egenskapen $XA = AX$.

6. Om A är en kvadratisk matris så definieras *spåret* av A (betecknas $\mathrm{tr}(A)$ eller $\mathrm{sp}(A)$) som summan av (huvud)diagonalelementen i A.

   (a) Visa att $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ och $\mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}(A)$, ($\lambda \in \mathbf{R}$).

   (b) Bevisa att om B är av typen m/n och C av typen n/m så är $\mathrm{tr}(BC) = \mathrm{tr}(CB)$.

7. Visa att om A är en godtycklig matris så är AA^T och A^TA symmetriska (en matris M är *symmetrisk* om $M^T = M$). Visa med exempel att dessa två produkter inte behöver vara lika. Visa också att $A + A^T$ är symmetrisk, om A är kvadratisk. Hur är det med $A - A^T$?

8. (a) Visa att det inte finns några 2/2-matriser A och B sådana att $AB - BA = I$ genom att studera det ekvationssystem som matrisekvationen definierar ($A - B$ betyder $A + (-1)B$).

   (b) Visa att ekvationen $AB - BA = I$ är omöjlig för vilka n/n-matriser A och B som helst. Ledning: Använd uppgift 6.

9. Bevisa att om matrisekvationen $AX = B$ har fler än en lösning, så har den oändligt många. Ledning: Om X_1 och X_2 är lösningar, undersök $X_3 = sX_1 + tX_2$ $(s, t \in \mathbf{R})$.
10. En kvadratisk matris sägs vara *idempotent*, om $A^2 = A$.

    (a) Visa att
$$A = \begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix} \quad \text{är idempotent.}$$

    (b) Visa att om $AB = A$ och $BA = B$, så är A och B idempotenta.
11. Sök en matris A som satisfierar ekvationen
$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} A \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
12. Undersök om det finns någon matris A som satisfierar ekvationen
$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 3 \end{pmatrix} = A \begin{pmatrix} 1 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix}.$$
13. En kvadratisk matris $A = (a_{ik})$ sägs vara *stokastisk* om $a_{ik} \geq 0$, för varje i och k, och om alla kolonnsummor blir 1 (dvs. $\sum_i a_{ik} = 1$, $k = 1, 2, \ldots$). Visa att produkten av två stokastiska matriser är en stokastisk matris.
14. Visa att om A är en idempotent matris, så är också följande matriser idempotenta
$$(a) \quad A^T, \qquad (b) \quad I - A, \qquad (c) \quad A^n, \quad n \in \mathbf{Z}_+.$$
15. Låt $A = (a_{ik})$, $B = (b_{ik})$ och $C = (c_{ik})$ beteckna m/n-matriser.

    (a) Antag först att $\mathbf{y}^T C\mathbf{x} = 0$ för alla kolonnvektorer $\mathbf{y}$ och $\mathbf{x}$ av typerna $m/1$ respektive $n/1$. Visa att $\mathbf{e}_i^T C\mathbf{e}_k = c_{ik}$, där $\mathbf{e}_j = (\ldots \; 0 \; 1 \; 0 \; \ldots)^T$ har en etta som j:te komponent men i övrigt består av nollor. Dra slutsatsen att $C = 0$.

    (b) Visa med hjälp av (a) att om $\mathbf{y}^T A\mathbf{x} = \mathbf{y}^T B\mathbf{x}$ för alla sådana $\mathbf{y}$ och $\mathbf{x}$ som i (a), så är $A = B$.

    (c) Visa med hjälp av (b) att om $A\mathbf{x} = B\mathbf{x}$ för varje $n/1$-vektor $\mathbf{x}$, så är $A = B$.
16. Lös matrisekvationen
$$\begin{pmatrix} a + 2b & 3a - b \\ c + 2d & 3c - d \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}.$$
17. Proteiner (P), kolhydrater (K) och fetter (F) bör finnas i en viss foderblandning i förhållandena 33 : 45 : 3. Dessa förekommer (mätt i gram per 100 g) i fettfri mjölk (M), sojamjöl (S) och vassla (V) enligt tabellen

	(M)	(S)	(V)
(P)	36	51	13
(K)	52	34	74
(F)	0	7	1

.

Bestäm (om möjligt) en blandning som ger de rätta proportionerna. (Använd decimaltal i kalkylerna.)

18. Låt $\mathbf{x}$ vara en $1/n$-radvektor och B en n/p-matris. Visa att matrisprodukten $\mathbf{x}B$ kan skrivas som en linjärkombination av raderna i B.

19. Låt A vara en m/n-matris sådan att AA^T är en m/m-nollmatris. Visa att $A = 0$.

20. Ställ upp ett ekvationssystem, vilkets lösning anger hur många molekyler x_i av varje slag som bör skrivas in i den kemiska formeln för att antalet atomer av de olika grundämnena skall vara detsamma både före och efter reaktionen:

$$(a) \quad \mathrm{x_1\,MnS + x_2\,As_2Cr_{10}O_{35} + x_3\,H_2SO_4 \rightarrow}$$
$$\mathrm{x_4\,HMnO_4 + x_5\,AsH_3 + x_6\,CrS_3O_{12} + x_7\,H_2O}\,,$$
$$(b) \quad \mathrm{x_1\,PbN_6 + x_2\,CrMn_2O_8 \rightarrow x_3\,Pb_3O_4 + x_4\,Cr_2O_3 + x_5\,MnO_2 + x_6\,NO}\,.$$
$$(c) \quad \mathrm{x_1\,NaHCO_3 + x_2\,H_3C_6H_5O_7 \rightarrow x_3\,Na_3C_6H_5O_7 + x_4\,H_2O + x_5\,CO_2}\,.$$

21. Betrakta en ekonomi bestående av tre sektorer: Kemikalier & Metaller (K), Energi (E) och Maskiner (M). Antag att varje sektor säljer en viss procent av sin produktion till de två andra enligt schemet

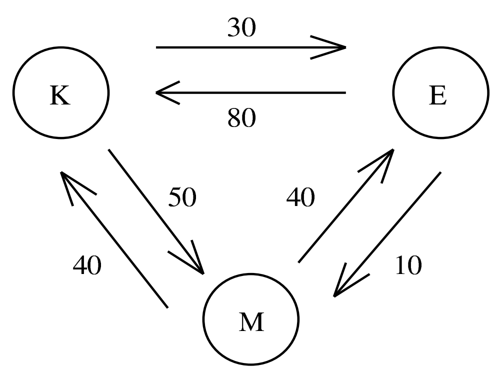

fig. 1

och använder resten själv. Bestäm priset på varje sektors produktion (per tidsenhet) så att ekonomin är i balans (varje sektors utgifter är lika stora som sektorns inkomster). Bortse från andra inkomster och utgifter än de som härrör från ovannämnda handel.

22. En bjälke, som stöds under bägge ändorna, påverkas av tre krafter f_1, f_2, f_3 som i fig. 2, varvid bjälken förskjuts nedåt sträckorna y_1, y_2, y_3. Låt $\mathbf{f}$ och $\mathbf{y}$ vara

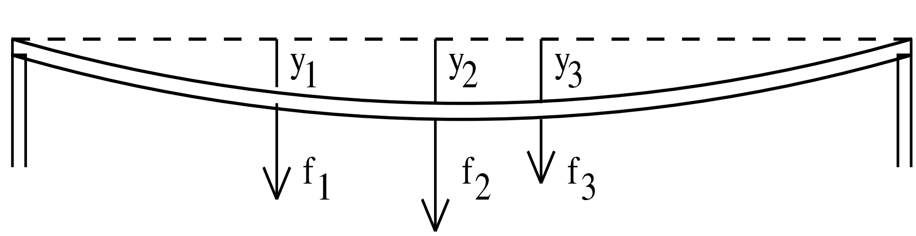

fig. 2

kolonnvektorerna med komponenterna f_i respektive y_i. Enligt Hookes lag är $\mathbf{y} = D\mathbf{f}$, där D är en 3/3-matris. Förklara hur kolonnerna i D (flexibilitetsmatrisen) kan bestämmas (mätas) genom att applicera lämpliga krafter. Vilken fysikalisk tolkning har kolonnerna i D^{-1} (styvhetsmatrisen)?

23. (*a*) I fig. 3 (a) är tvärsnittet genom en metallbjälke uppritad. Den vänstra, högra, övre respektive undre sidan av bjälken har av yttre orsaker temperaturen 10°C, 40°C, 30°C respektive 20°C grader. Eftersom termisk jämvikt antas råda, är temperaturen T_k i varje nod (noderna är utmärkta med cirklar i figurerna) approximativt medeltalet av temperaturerna i de angränsande noderna (t.ex. är $4T_1 = 10 + 30 + T_2 + T_3$). Bestäm temperaturen i varje nod.

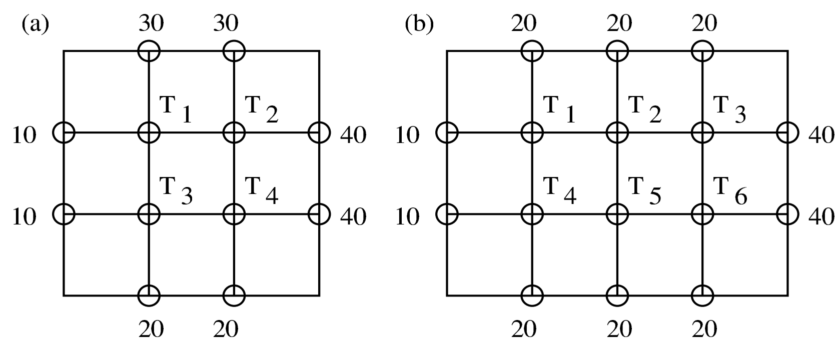

fig. 3

(*b*) Gör detsamma för tvärsnittet i fig. 3 (b). Ledning: Utnyttja symmetrin i temperaturfördelningen för att förenkla kalkylerna.

3. Gausseliminering genom matrismultiplikation

Vi skall visa att den omformning, som man får genom att utföra en basoperation på ett räkneschema, också kan fås genom att man multiplicerar räkneschemat från vänster med en viss matris E. Detta är ekvivalent med att man transformerar motsvarande ekvation $A\mathbf{x} = \mathbf{b}$ genom att multiplicera den från vänster med E, vilket ger $EA\mathbf{x} = E\mathbf{b}$.

Låt oss först se hur detta fungerar för $(BO1)$:

Exempel 3.1. Betrakta ekvationssystemet

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 6 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 4x_2 - 6x_3 &= 1\,, \end{aligned}$$

som vi skriver som ett räkneschema $(A|\mathbf{b})$ genom att sätta

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \\ 1 & 4 & -6 \end{pmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 1 \end{pmatrix}.$$

Det första elimineringssteget, $(2) \to (2) - 2\,(1)$, är

$$(A|\mathbf{b}) = \begin{pmatrix} 1 & 2 & 3 & | & 6 \\ 2 & 5 & 1 & | & 9 \\ 1 & 4 & -6 & | & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & | & 6 \\ 0 & 1 & -5 & | & -3 \\ 1 & 4 & -6 & | & 1 \end{pmatrix} = (U_1|\mathbf{b}_1)\,.$$

Sätt

$$E_1 = I - 2J_{21} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

där J_{21} är en 3/3-matris som har nollor på alla andra platser än $(2,1)$, där den har en etta. Då svarar multiplikation med E_1 mot det första elimineringssteget, ty

$$E_1(A|\mathbf{b}) = (E_1A|E_1\mathbf{b})\,,$$

där enligt distributionslagen

$$E_1A = (I - 2J_{21})A = A - 2J_{21}A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \\ 1 & 4 & -6 \end{pmatrix} - 2\begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} = U_1$$

och på samma sätt

$$E_1\mathbf{b} = \mathbf{b} - 2J_{21}\mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 0 \\ 6 \\ 0 \end{pmatrix} = \mathbf{b}_1\,.$$

Det andra elimineringssteget $(3) \to (3) - 1\,(1)$ svarar enligt samma mönster mot multiplikation från vänster med $E_2 = I - 1J_{31}$,

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \text{ som ger } (U_2|\mathbf{b}_2) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & -5 & -3 \\ 0 & 2 & -9 & -5 \end{array}\right).$$

Det tredje, $(3) \to (3) - 2\,(2)$, svarar i sin tur mot multiplikation med $E_3 = I - 2J_{32}$,

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, \text{ och ger } (U_3|\mathbf{b}_3) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Systemet $A\mathbf{x} = \mathbf{b}$ har nu överförts på echelonform genom multiplikation från vänster med $E = E_3E_2E_1$.

Allmänt gäller för ett system $A\mathbf{x} = \mathbf{b}$, där A är en m/n-matris, att $(BO1)$-operationen

$$\{(i) \to (i) - \lambda(k)\} \quad \text{motsvarar} \quad \left\{ \begin{array}{c} \text{multiplikation fr.v. med} \\ E = I - \lambda J_{ik} \end{array} \right\},$$

där J_{ik} är en m/m-matris som består av bara nollor utom på platsen (i, k) där den har en etta.

För $(BO2)$ har vi sambandet: Operationen

$$(i) \leftrightarrow (k)$$

motsvarar multiplikation från vänster med

$$(1) \qquad E = \begin{pmatrix} \ddots &&&&&&&& \\ & 1 &&&&&&& \\ && 0 & \cdots &&& 1 && \\ && \vdots & 1 &&& \vdots && \\ && \vdots && \ddots && \vdots && \\ && \vdots &&& 1 & \vdots && \\ && 1 & \cdots &&& 0 && \\ &&&&&&& 1 & \\ &&&&&&&& \ddots \end{pmatrix} \begin{matrix} \\ \\ (i) \\ \\ \\ \\ (k) \\ \\ \\ \end{matrix},$$

där (enligt konventionen) ett tomrum står för en nolla. Matrisen E är i detta fall en s.k. *enkel permutationsmatris*, som fås ur en enhetsmatris I så att man i denna låter raderna (i) och (k) byta plats.

För $(BO3)$ gäller: Operationen

$$(i) \to \lambda\,(i)$$

svarar mot multiplikation från vänster med diagonalmatrisen

$$E = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \begin{matrix} \\ \\ \\ (i) \\ \\ \\ \\ \end{matrix}.$$

Exempel 3.2. För tvåradiga system har vi t.ex.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right) = \left(\begin{array}{cc|c} c & d & f \\ a & b & e \end{array}\right), \qquad (BO2),$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 7 \end{pmatrix} \left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right) = \left(\begin{array}{cc|c} a & b & e \\ 7c & 7d & 7f \end{array}\right), \qquad (BO3).$$

Det faktum, att $(BO1)$–$(BO3)$ svarar mot multiplikation (fr.v.) med vissa matriser, kommer att tillåta oss att dra långtgående slutsatser. För att kunna dra dessa slutsatser, måste vi emellertid utveckla mera grundteori.

Matrisinverser

För en basoperation gäller alltid (se Sats 1.1) att en lämplig annan basoperation kan upphäva den effekt som den första basoperationen hade. Om någon basoperation således transformerar systemet $A\mathbf{x} = \mathbf{b}$ till $EA\mathbf{x} = E\mathbf{b}$, så kan dess verkan upphävas genom multiplikation från vänster med en ny matris E', som svarar mot den "inversa" basoperationen:

$$E'EA\mathbf{x} = E'E\mathbf{b}.$$

Detta system bör nu vara identiskt med det ursprungliga, så att $E'E = I$, vilket leder till följande definition:

Definition 3.1. Två n/n-matriser A och B är varandras *inverser* om

$$AB = I \quad \text{och} \quad BA = I.$$

En kvadratisk matris som har en invers sägs vara *inverterbar*.

Exempel 3.3. Vi konstaterar först att det finns matriser, som är olika nollmatrisen men som saknar invers. Betrakta t.ex. matrisen

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

För varje $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ är nämligen

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix} \neq I,$$

dvs. ingen matris B är invers till A.

Sats 3.1. *En kvadratisk matris A har högst en invers.*

Bevis. Om B och B' är inverser till A så är

$$BA = I \quad \text{och} \quad AB' = I.$$

Alltså är $B = BI = B(AB') = (BA)B' = IB' = B'$. $\diamondsuit$

För inversen till en matris A skall vi från och med nu använda beteckningen A^{-1}. Observera att om $B = A^{-1}$ så är enligt definitionen $A = B^{-1}$, vilket genom insättning ger att $(A^{-1})^{-1} = A$. Detta är den första av räknereglerna i följande sats:

Sats 3.2. *Om A är en inverterbar n/n-matris så är A^{-1} inverterbar och λA är inverterbar om dessutom $\lambda \neq 0$. Under de nämnda antagandena är*

$$(i) \quad (A^{-1})^{-1} = A\,,$$
$$(ii) \quad (\lambda A)^{-1} = \frac{1}{\lambda} A^{-1} \quad (\lambda \neq 0)\,.$$

Om både A och B är inverterbara n/n-matriser, så är AB inverterbar och

$$(iii) \quad (AB)^{-1} = B^{-1}A^{-1}\,.$$

Bevis. Vi bevisade redan (i). För att bevisa (ii) räcker det, på grund av Sats 3.1, att visa att $\frac{1}{\lambda}A^{-1}$ är en invers till λA: Eftersom skalära faktorer kan placeras var som helst i en matrisprodukt är

$$(\lambda A)\left(\frac{1}{\lambda}A^{-1}\right) = \left(\lambda\frac{1}{\lambda}\right) AA^{-1} = I$$
$$\left(\frac{1}{\lambda}A^{-1}\right)(\lambda A) = \left(\frac{1}{\lambda}\lambda\right) A^{-1}A = I\,.$$

För verifieringen av (iii) räcker det att visa att $B^{-1}A^{-1}$ är en invers till AB:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I\,,$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I\,. \diamond$$

Vi har sett att $(BO1)$ svarar mot multiplikation fr.v. med en matris av formen $E = I - \lambda J_{ik}$ $(i \neq k)$. Eftersom J_{ik} består av nollor utom på platsen (i, k), där den har en etta, inses lätt att

$$J_{ik}J_{lm} = \begin{cases} J_{im}, & \text{om } k = l, \\ 0, & \text{om } k \neq l. \end{cases} \tag{2}$$

På grund av detta är $E^{-1} = I + \lambda J_{ik}$, ty vi har t.ex. att

$$\begin{aligned} EE^{-1} &= (I - \lambda J_{ik})(I + \lambda J_{ik}) \\ &= I + \lambda J_{ik} - \lambda J_{ik} - \lambda^2 J_{ik}J_{ik} = I\,. \end{aligned}$$

På samma sätt ser man att $E^{-1}E = I$.

Exempel 3.4. Om $E = I - 5J_{12}$ är en 3/3-matris, är

$$E = \begin{pmatrix} 1 & -5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{och} \quad E^{-1} = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Detta betyder att om man utför $(BO1)$-operationerna $(1) \to (1) - 5\,(2)$ och $(1) \to (1) + 5\,(2)$ efter varandra, så förblir räkneschemat (eller ekvationssystemet) oförändrat.

Exempel 3.5. Den enkla permutationsmatrisen $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ är sin egen invers, ty

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

Detta svarar mot att om man utför $(BO2)$-operationen $(1) \leftrightarrow (2)$ två gånger på ett räkneschema, så återfår man det ursprungliga räkneschemat.

Exempel 3.6. Matrisen

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{har inversen} \qquad E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

vilket är liktydigt med att $(BO3)$-operationerna $(2) \to 3\,(2)$ och $(2) \to \frac{1}{3}(2)$ upphäver varandra.

Uträkning av inverser

Att bestämma inversen $X = A^{-1}$ till en n/n-matris A innebär att man löser matrisekvationen $AX = I$. Vi kommer nämligen senare att se, att om en matris X satisfierar denna ekvation så satisfierar samma X **automatiskt** också ekvationen $XA = I$. På bl.a. detta grundar sig följande metod att räkna ut inversen till en matris A:

Vi inför beteckningar för kolonnerna i X och I genom att sätta $X = (\mathbf{x}_1 \cdots \mathbf{x}_n)$ och $I = (\mathbf{e}_1 \cdots \mathbf{e}_n)$. Matrisekvationen $AX = I$ kan då skrivas som n ekvationssystem med samma koefficientmatris A:

$$A\mathbf{x}_1 = \mathbf{e}_1, \quad A\mathbf{x}_2 = \mathbf{e}_2, \quad \ldots \quad , A\mathbf{x}_n = \mathbf{e}_n\,.$$

Dessa kan lösas var och en för sig med hjälp av räknescheman $(A|\mathbf{e}_1), \ldots, (A|\mathbf{e}_n)$. Men då det ju är koefficientmatrisen A, som bestämmer vilka basoperationer man använder, kommer man att upprepa (nästan) samma kalkyl n gånger. För att undvika dessa upprepningar använder man sig av **ett enda** sammanslaget räkneschema

$$(A|\mathbf{e}_1 \ldots \mathbf{e}_n) = (A|I)\,,$$

där alla högerled beaktas samtidigt genom att man skriver dem efter varandra. Om en lösning X existerar, kan detta schema fås i reducerad echelonform med hjälp av basoperationer enligt mönstret

$$(A|I) \to \cdots \to (I|\mathbf{x}_1 \cdots \mathbf{x}_n) = (I|X)\,.$$

Lösningen $\mathbf{x}_i$ till det i:te ekvationssystemet finner man då på den plats där högerledet $\mathbf{b}_i$ ursprungligen fanns. Detta innebär att då det i räkneschemat står en enhetsmatris I på vänstra sidan om det vertikala strecket, står inversen $X = A^{-1}$ på den högra sidan om strecket.

Exempel 3.7. Vi inverterar matrisen

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

med hjälp av följande kalkyler

$$\left(\begin{array}{cc|cc} -2 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \end{array}\right) \overset{BO2}{\rightarrow} \left(\begin{array}{cc|cc} 1 & -2 & 0 & 1 \\ -2 & 1 & 1 & 0 \end{array}\right) \overset{BO1}{\rightarrow} \left(\begin{array}{cc|cc} 1 & -2 & 0 & 1 \\ 0 & -3 & 1 & 2 \end{array}\right)$$
$$\overset{BO3}{\rightarrow} \left(\begin{array}{cc|cc} 1 & -2 & 0 & 1 \\ 0 & 1 & -1/3 & -2/3 \end{array}\right) \overset{BO1}{\rightarrow} \left(\begin{array}{cc|cc} 1 & 0 & -2/3 & -1/3 \\ 0 & 1 & -1/3 & -2/3 \end{array}\right).$$

Inversen är alltså

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

Exempel 3.8. Matrisen $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ har ingen invers, ty den andra raden blir inkonsistent:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right).$$

LU-faktorisering

I exempel 3.1 har vi genom tillämpning av ($BO1^+$) fått ett system på echelonform. Detta svarade mot att ekvationen $A\mathbf{x} = \mathbf{b}$ multiplicerades från vänster med $E = E_3E_2E_1$ så att en ny ekvation $U_3\mathbf{x} = \mathbf{b}_3$ uppstod, där U_3 är en echelonmatris. Nu är ju $EA = U_3$, vilket ger, då vi multiplicerar från vänster med $L = E^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}$:

$$A = LU_3\,.$$

Hur ser L ut? Inverserna av E_1, E_2 och E_3 fås på samma sätt som i exempel 3.4:

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

Observera att elementen på platserna $(2,1)$, $(3,1)$ och $(3,2)$ i L är just precis de multipler $\lambda = 2$, 1 och 2 som vi använde i de tre elimineringsstegen då vi **subtraherade** en multipel av en ekvation från en annan. Detta är ingen slump utan vi kan formulera en generell regel. För att förtydliga detta skall vi se närmare på hur produkten L i vårt exempel uppstod: På grund av den speciella ordning, i vilken vi eliminerat (först i kolonn 1, sedan i kolonn 2 osv.), blir enligt formel (2) produkter av minst två faktorer J_{ik} en nollmatris, så att

$$L = (I + 2J_{21})(I + 1J_{31})(I + 2J_{32}) = I + 2J_{21} + 1J_{31} + 2J_{32}\,.$$

Detta betyder att matrisen L kan skrivas ut utan att man utför någon multiplikation. Talet λ hamnar alltid på platsen (i,k) i L då man utför operationen $(i) \to (i) - \lambda(k)$:

Sats 3.3. *Om en m/n-matris A kan överföras på echelonform U enbart med hjälp av $(BO1^+)$, så kan A LU-*faktoriseras *i en produkt*

$$A = LU\,.$$

Om elimineringen skett i normal ordning så att man eliminerat i A:s kolonner från vänster till höger, kan matrisen L skrivas ut enligt regeln:

L är en nedåt triangulär m/m-matris, sådan att

- *L har ettor i diagonalen;*
- *multipeln λ sätts på plats (i,k) under diagonalen i L om man eliminerat med operationen $(i) \to (i) - \lambda(k)$;*
- *övriga matriselement är nollor.*

Om A är en n/n-matris och antalet pivotelement i U är precis n, så är U en uppåt triangulär matris, vilkens diagonalelement (= pivotelement) *alla är olika noll.*

Vid LU-faktorisering av en matris A speler högerledet ingen roll. Vi skriver därför inte ut något högerled i räkneschemat.

Exempel 3.9. Om man skall LU-faktorisera matrisen

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 4 & 5 \end{pmatrix}$$

gör man följande kalkyler med hjälp av $(BO1^+)$:

$$A \underset{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}{\longrightarrow} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix} \underset{(3)}{\to} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = U\,,$$

där de multipler λ som använts i operationerna antecknats under pilarna. T.ex. operationen $(3) \to (3) - 3\,(2)$ ger upphov till talet 3 på plats $(3,2)$ i L:

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

Om man numrerar pivotkolonnerna i A (= de kolonner som innehåller pivotelement) från vänster till höger, kan man också uttrycka regeln i Sats 3.3 på följande sätt: **Då man eliminerar i** *pivotkolonn* **nummer k i A, skall de använda multiplerna λ skrivas i** *kolonn* **nummer k i L.**

Permutationsmatriser

Ibland kan en matris inte fås i echelonform enbart med hjälp av $(BO1^+)$. Den kan då inte LU-faktoriseras **direkt.** Detta gäller exempelvis för matrisen

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

Ett ekvationssystem som svarar mot A är t.ex.

$$\begin{aligned} 2x_2 &= b_1 \\ 3x_1 + 4x_2 &= b_2 \,. \end{aligned}$$

Om man permuterar ekvationerna, dvs. använder $(BO2)$, blir systemet däremot genast uppåt triangulärt. Koefficientmatrisen för det nya systemet blir

$$\begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}$$

och fås ur A genom multiplikation från vänster med en permutationsmatris P av det slag som beskrivs i ekvation (1).

Definition 3.2. En matris, som fås ur en enhetsmatris I genom att man låter raderna i I byta ordning, kallas en *permutationsmatris.* En permutationsmatris är *enkel* och betecknas med P_{ik} om den fås ur I genom att man låter raderna (i) och (k) i I byta plats $(i \neq k)$.

Notera att $P_{ik} = P_{ki}$ och att P_{ik} är inverterbar och är sin egen invers: $P_{ik}^{-1} = P_{ik}$.

Vi har sett att multiplikation av en matris A från vänster med den enkla permutationsmatrisen P_{ik} åstadkommer att raderna (i) och (k) i A byter plats. Något liknande gäller för alla permutationsmatriser:

Exempel 3.10. Betrakta t.ex. permutationsmatrisen

$$P = \begin{pmatrix} \bar{\mathbf{e}}_2 \\ \bar{\mathbf{e}}_3 \\ \bar{\mathbf{e}}_1 \\ \bar{\mathbf{e}}_4 \end{pmatrix},$$

där $\bar{\mathbf{e}}_i$ betecknar rad i i enhetsmatrisen I_4. Ordningsföljden på raderna i P kan fås ur den ursprungliga (den som de har i I_4) med hjälp av en följd av enkla permutationer:

$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \; \xrightarrow{P_{12}} \; \begin{matrix} 2 \\ 1 \\ 3 \\ 4 \end{matrix} \; \xrightarrow{P_{23}} \; \begin{matrix} 2 \\ 3 \\ 1 \\ 4 \end{matrix}.$$

Således är $P = P_{23}P_{12}$. Om vi nu bildar en produkt PA så kommer P_{12} att byta om raderna (1) och (2) i A, varefter P_{23} byter om raderna (2) och (3) i $P_{12}A$. Resultatet blir att P flyttar om raderna i A så att de får ordningsföljden (2), (3), (1), (4).

Eftersom enkla permutationsmatriser är sina egna inverser, gäller dessutom att P är inverterbar och att $P^{-1} = P_{12}P_{23}$.

Detta gäller generellt:

Sats 3.4. *Låt P vara den permutationsmatris som fås ur I_m genom att man ordnar om raderna i ordningsföljden σ samt låt A vara en m/n-matris. Då är*

(i) P en produkt $P = P_{i_1j_1} \cdots P_{i_pj_p}$ av enkla permutationsmatriser;

(ii) P inverterbar och $P^{-1} = P_{i_pj_p} \cdots P_{i_1j_1}$;

(iii) PA är likamed den matris som man får ur A genom att skriva A:s rader i ordningsföljden σ.

Mera om LU-faktorisering

Vi har sett att en matris kan LU-faktoriseras om den kan överföras i echelonform enbart med hjälp av $(BO1^+)$. Detta är inte alltid möjligt utan man kan bli tvungen att permutera rader med hjälp av $(BO2)$ för att komma till en echelonform. För att inkludera också sådana fall krävs en ny formulering av satsen om LU-faktorisering.

Antag att A är en m/n-matris, som (eventuellt) inte kan fås på echelonform enbart med hjälp av $(BO1^+)$ så att också $(BO2)$ måste tas till hjälp. Då kan man tänka sig att **genast i början** göra de radbyten som kommer att behövas (dvs. bilda en matris PA där P är en permutationsmatris) för att i fortsättningen uteslutande använda $(BO1^+)$, varvid man LU-faktoriserar PA. Detta är möjligt för alla m/n-matriser A, ty $(BO3)$ behövs strängt taget inte vid omformning till echelonform. Däremot behövs $(BO3)$ nog om man vill få fram en reducerad echelonform.

Vi kan alltså formulera följande sats:

Sats 3.5. *För varje m/n-matris A finns en permutationsmatris P, sådan att PA kan LU-faktoriseras*

$$PA = LU\,.$$

I praktiken behöver man inte göra alla radbyten först, utan man kan under räkningens gång modifiera sina listor med multipler λ efter varje radbyte, så att listorna gäller **som om radbytet hade gjorts redan i början**:

Exempel 3.11. I följande sekvens av räknescheman skriver vi först ut en preliminär lista med multipler (nedan med index 0), som byts ut mot en ny lista (med index 1) så snart operation $(BO2)$ används. En multipel hör ju ihop med en viss rad, varför multiplerna bör byta plats på samma sätt som motsvarande rader:

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & 1 & 1 \\ 1 & 3 & 2 & 2 \end{pmatrix} \underset{\begin{pmatrix} 2 \\ 1 \end{pmatrix}_0}{\xrightarrow{BO1^+}} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 1 & 3 & 2 \end{pmatrix}$$

$$\underset{\begin{pmatrix} 1 \\ 2 \end{pmatrix}_1}{\xrightarrow{BO2}} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & 1 \end{pmatrix} \begin{matrix} (1) \\ (3) \\ (2) \end{matrix} = U\,.$$

Samtidigt har vi noterat att den ursprungliga ordningen (1), (2), (3) på raderna i A har omvandlats till (1), (3), (2) i U, vilket ger oss permutationsmatrisen P. Vi har alltså att $PA = LU$, där

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{och} \quad L = \begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & 0 & 1 \end{pmatrix} .$$

Definition 3.3. En n/n-matris A sägs vara *icke-singulär* om den med hjälp av (vilka som helst) basoperationer kan fås i en uppåt triangulär form U, som innehåller n pivotelement. Om antalet pivotelement är mindre än n sägs matrisen A vara *singulär*.

Exempel 3.12. Matrisen

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 2 & 8 & 8 \end{pmatrix}$$

är icke-singulär, ty echelonformen nedan innehåller tre pivotelement:

$$A \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{pmatrix} .$$

Däremot är matrisen

$$B = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 3 & 9 & 8 \end{pmatrix}$$

singulär, ty

$$B \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{pmatrix} .$$

Nedanstående ekvivalenser för en n/n-matris A med den motsvarande echelonformen U,

$$\begin{aligned} A \text{ är singulär} &\iff \#(\text{ pivotelement}) < n \\ &\iff \text{ekvationen } U\mathbf{x} = \mathbf{0} \text{ har icke-triviala lösningar} \\ &\iff \text{ekvationen } A\mathbf{x} = \mathbf{0} \text{ har icke-triviala lösningar} \ , \end{aligned}$$

ger en enkel karakterisering av en icke-singulär matris, vilken vi i fortsättningen kommer att behöva i många sammanhang:

$$(3) \qquad A \text{ är icke-singulär} \quad \iff \quad \left\{ \begin{array}{c} A \text{ är kvadratisk och} \\ A\mathbf{x} = \mathbf{0} \text{ bara om } \mathbf{x} = \mathbf{0} \end{array} \right\} .$$

Vi kan nu göra ett tillägg till Sats 3.5:

Sats 3.6. *Antag att A är kvadratisk och att $PA = LU$ är LU-faktoriseringen av PA, där P är en permutationsmatris. Om A är* icke-singulär, *så gäller:*

(*i*) *Diagonalelementen i U är olika noll;*

(*ii*) *ekvationen $A\mathbf{x} = \mathbf{b}$ har* exakt en *lösning (för varje $\mathbf{b}$).*

Om A är singulär *så är något av diagonalelementen i U en nolla.*

Bevis. Det sista påståendet och (*i*) följer direkt ur definition 3.3. Ekvationen $A\mathbf{x} = \mathbf{b}$ är ekvivalent med en ekvation av formen $U\mathbf{x} = \mathbf{b}_1$. Eftersom varje kolonn i U innehåller ett pivotelement om A är icke-singulär, så är ingen variabel fri och systemet är konsistent. Bakåtsubstitution ger därför exakt en lösning. Alltså gäller (*ii*). $\diamondsuit$

LU-faktoriseringen av en kvadratisk matris A är osymmetrisk såtillvida att L har ettor i diagonalen medan U inte behöver ha det som t.ex. i LU-faktoriseringen

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

Detta skönhetsfel är lätt att korrigera. Vi skriver U som en produkt av en diagonalmatris D och en ny uppåt triangulär matris U' med ettor i diagonalen,

$$U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 1/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = DU'\,,$$

och får $A = LDU'$. En sådan LDU-faktorisering av A (eller PA) är alltid möjlig om diagonalelementen i U är olika noll, dvs. om A är icke-singulär.

Sats 3.7. *Om A är en icke-singulär kvadratisk matris så finns det en permutationsmatris P och en så kallad LDU*-faktorisering *$PA = LDU$ av PA, där*

- *L är en nedåt triangulär matris med ettor i diagonalen,*
- *D är en diagonalmatris, vilkens diagonalelement är olika noll,*
- *U är en uppåt triangulär matris med ettor i diagonalen.*

Användningar av LU- och LDU-faktoriseringar

Om man samtidigt löser flera ekvationssystem,

$$A\mathbf{x}_1 = \mathbf{b}_1, \quad A\mathbf{x}_2 = \mathbf{b}_2, \quad \dots \quad , A\mathbf{x}_p = \mathbf{b}_p\,,$$

med samma koefficientmatris, kan man göra som vi gjorde vid invertering av en matris: I stället för att göra skilda kalkyler i flera räknescheman $(A|\mathbf{b}_1), \dots, (A|\mathbf{b}_p)$, vilket leder till upprepningar av räkneoperationer, slår vi ihop dessa till ett enda räkneschema

$$(A|B) = (A|\mathbf{b}_1 \dots \mathbf{b}_p)\,,$$

genom att skriva högerleden efter varandra i $B = (\mathbf{b}_1 \dots \mathbf{b}_p)$. Lösningarna till alla dessa system kan sedan lätt skrivas ut då schemat $(A|B)$ överförts i reducerad echelonform. **Observera att detta är ekvivalent med att lösa matrisekvationen $AX = B$, där $X = (\mathbf{x}_1 \dots \mathbf{x}_p)$** (se uppgift 3 i kapitel 2).

Ovanbeskrivna metod, som är bra vid räkning för hand, har olägenheten att alla högerled måste vara kända innan man kan påbörja kalkylerna.

En annan metod använder sig av LU-faktoriseringen $A = LU$ av A (eller av $PA = LU$ av PA, varvid högerleden är $P\mathbf{b}_1, \dots, P\mathbf{b}_p$). Ekvationen $A\mathbf{x}_i = \mathbf{b}_i$ kan skrivas $LU\mathbf{x}_i = \mathbf{b}_i$, och denna i sin tur som två ekvationer

$$\begin{aligned} L\mathbf{y}_i &= \mathbf{b}_i \\ U\mathbf{x}_i &= \mathbf{y}_i\,, \end{aligned} \qquad (i = 1, \dots, p) \tag{4}$$

där man har infört en temporär variabel $\mathbf{y}_i$. Först löser man $A\mathbf{x}_1 = \mathbf{b}_1$ och får LU-faktoriseringen på köpet. Efter det använder man sig av (4), för $i = 2, \dots, p$, så att man löser den första ekvationen med räkneschemat $(L|\mathbf{b}_i)$, vilket ger $\mathbf{y}_i$, varefter man löser den andra ekvationen med räkneschemat $(U|\mathbf{y}_i)$, vilket ger $\mathbf{x}_i$. Det är uppenbart att **alla högerled inte behöver vara kända från början** då man använder denna metod, utan nya högerled kan läggas till när som helst (se uppgifterna 3 och 10).

På grund av att både L och U är triangulära, krävs bara framåt- respektive bakåtsubstitutioner vid lösningen, dvs. ett relativt litet antal räkneoperationer. De båda beskrivna metoderna kräver ungefär lika stor arbetsinsats. Man kan visa att om A är en n/n-matris så krävs i bägge fallen

$$\approx \frac{2}{3}n^3 + 2pn^2$$

räkneoperationer med de fyra räknesätten.

Exempel 3.13. Om vi skall lösa systemen $A\mathbf{u} = \mathbf{b}_1$ och $A\mathbf{v} = \mathbf{b}_2$, där

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & 1 \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{och} \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

ställer vi enligt den först beskrivna metoden upp ett gemensamt räkneschema, som överförs i reducerad echelonform:

$$(A|\mathbf{b}_1\mathbf{b}_2) = \left(\begin{array}{ccc|cc} 1 & 2 & -1 & 1 & -1 \\ 1 & 3 & 1 & 2 & 2 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|cc} 1 & 0 & -5 & -1 & -7 \\ 0 & 1 & 2 & 1 & 3 \end{array}\right).$$

Både u_3 och v_3 är fria, så vi kan sätta $u_3 = s$ och $v_3 = t$. Med hjälp av den första kolonnen i högerledet får vi lösningarna till det första systemet och med hjälp av den andra kolonnen i högerledet lösningarna till det andra systemet:

$$\begin{aligned} u_1 &= -1 + 5s \\ u_2 &= \ \ 1 - 2s \\ u_3 &= \qquad\ \ s \end{aligned} \qquad \text{och} \qquad \begin{aligned} v_1 &= -7 + 5t \\ v_2 &= \ \ 3 - 2t\,. \\ v_3 &= \qquad\ \ t \end{aligned}$$

Enligt den andra metoden ställer vi upp räkneschemat för den första ekvationen och kan, då vi har nått echelonform (med hjälp av $(BO1^+)$ och eventuellt $(BO2)$)

$$\begin{pmatrix} 1 & 2 & -1 & | & 1 \\ 1 & 3 & 1 & | & 2 \end{pmatrix} \underset{(1)}{\rightarrow} \begin{pmatrix} 1 & 2 & -1 & | & 1 \\ 0 & 1 & 2 & | & 1 \end{pmatrix},$$

skriva ut LU-faktoriseringen

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

Sedan fortsätter vi tills räkneschemat är i reducerad echelonform och kan då skriva ut lösningarna till det första systemet. För att lösa det andra systemet betraktar vi först $L\mathbf{y} = \mathbf{b}_2$,

$$(L|\mathbf{b}_2) = \begin{pmatrix} 1 & 0 & | & -1 \\ 1 & 1 & | & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & | & -1 \\ 0 & 1 & | & 3 \end{pmatrix}, \text{ och får att } \mathbf{y} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Därefter löser vi $U\mathbf{v} = \mathbf{y}$,

$$\begin{pmatrix} 1 & 2 & -1 & | & -1 \\ 0 & 1 & 2 & | & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -5 & | & -7 \\ 0 & 1 & 2 & | & 3 \end{pmatrix},$$

och finner den tidigare nämnda lösningen.

Kom ihåg att vårt exempel med små matriser bara visar principen för hur metoderna fungerar. Fördelarna hos dessa två metoder blir betydande först då systemen är stora och/eller antalet system stort.

Vi skall nu se vilka **teoretiska** följder Sats 3.7 om LDU-faktorisering har. Vi visar först att om A är en icke-singulär n/n-matris med LDU-faktoriseringen $A = LDU$, så är varje faktor L, D och U inverterbar:

Faktorn L är en produkt $L = E_1^{-1} \cdots E_p^{-1}$ av inverterbara faktorer (som svarar mot $(BO1^+)$) och är därmed själv inverterbar enligt Sats 3.2 (*iii*). Faktorn

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \quad \text{har inversen} \quad D^{-1} = \begin{pmatrix} d_1^{-1} & & \\ & \ddots & \\ & & d_n^{-1} \end{pmatrix}.$$

Varje diagonalelement d_i är nämligen olik noll enligt Sats 3.6. Den uppåt triangulära matrisen U slutligen, kan överföras på en reducerad echelonform som är I med hjälp av bakåtsubstitutioner $(BO1^-)$, dvs. genom att man multiplicerar med inverterbara matriser F_j av formen $I - \lambda J_{ik}$:

$$F_q \cdots F_1 U = I.$$

Om man multiplicera från vänster med i tur och ordning $F_q^{-1}, \ldots, F_1^{-1}$, får man att U är en produkt av inverterbara faktorer, $U = F_1^{-1} \cdots F_q^{-1}$. Alltså är också U inverterbar.

Med hjälp av detta kan vi visa att LDU-faktoriseringar är **entydiga**:

Sats 3.8. *Antag att A är en icke-singulär kvadratisk matris och att A* (eller PA, där P är en permutationsmatris) *har LDU-faktoriseringarna*

$$A = L_1 D_1 U_1 \qquad \text{och} \qquad A = L_2 D_2 U_2 \,.$$

Då är $L_1 = L_2$, $D_1 = D_2$ och $U_1 = U_2$.

Bevis. Ur $L_1 D_1 U_1 = L_2 D_2 U_2$ följer, då man multiplicerar från vänster med L_1^{-1} och D_1^{-1} och från höger med U_2^{-1}, att

$$U_1 U_2^{-1} = D_1^{-1} L_1^{-1} L_2 D_2 \,.$$

Vänstra ledet är en produkt av två uppåt triangulära matriser med ettor i diagonalen och har därför själv samma form (se uppgift 19) medan högra ledet är en nedåt triangulär matris. Detta medför att bägge leden måste vara I:

$$U_1 U_2^{-1} = I \quad \text{och} \quad D_1^{-1} L_1^{-1} L_2 D_2 = I \,. \tag{5}$$

Multiplikation av den första likheten i (5) med U_2 från höger ger att $U_1 = U_2$. Genom att multiplicera den andra likheten från vänster med D_1 och från höger med D_2^{-1} fås

$$L_1^{-1} L_2 = D_1 D_2^{-1} \,,$$

där vänsterledet är nedåt triangulärt med ettor i diagonalen och högerledet en diagonalmatris. Bägge leden måste då vara en enhetsmatris: $L_1^{-1} L_2 = I$ och $D_1 D_2^{-1} = I$. Följaktligen är $L_1 = L_2$ och $D_1 = D_2$. $\diamondsuit$

Anmärkning. Observera att det i allmänhet finns flera olika permutationsmatriser P, sådana att PA kan LDU-faktoriseras för en given icke-singulär matris A. Dessa faktoriseringar är olika för olika P men då P är bestämt är faktorerna i LDU-faktoriseringen av PA entydiga.

En viktig konsekvens av inverterbarheten hos faktorerna L, D och U i en LDU-faktorisering är följande sats:

Sats 3.9. *En kvadratisk matris A är icke-singulär om och endast om den är inverterbar.*

Bevis. Antag att A är icke-singulär. Då finns det en permutationsmatris P och en tillhörande LDU-faktorisering $PA = LDU$. Eftersom permutationsmatriser är inverterbara (Sats 3.4), är

$$A = P^{-1} LDU \,,$$

där alla faktorer är inverterbara. Enligt Sats 3.2 (*iii*) är då också A inverterbar.

Antag omvänt att A är inverterbar. För att visa att A är icke-singulär, antar vi att $\mathbf{x}$ är en någon lösning till ekvationen $A\mathbf{x} = \mathbf{0}$. Genom multiplikation från vänster med A^{-1} fås att $\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$ är den enda lösningen. Således är A en icke-singulär matris enligt (3). $\diamondsuit$

Symmetriska matriser

Definition 3.4. En matris A är *symmetrisk* om $A^T = A$.

Om A är av typen m/n så är A^T av typen n/m. Således måste A vara kvadratisk för att kunna vara symmetrisk. Dessutom bör $a_{ik} = a_{ki}$ för varje i och k hos en symmetrisk matris.

Sats 3.10. *Om matrisen A är inverterbar, är också A^T inverterbar och*

$$(A^T)^{-1} = (A^{-1})^T .$$

Bevis. Om A är inverterbar, är

$$(A^T)(A^{-1})^T = (A^{-1}A)^T = I^T = I ,$$
$$(A^{-1})^T(A^T) = (AA^{-1})^T = I^T = I ,$$

dvs. $(A^{-1})^T$ är inversen till A^T. $\diamondsuit$

Ur Sats 3.10 följer:

Sats 3.11. *Om A är inverterbar och symmetrisk så är också A^{-1} symmetrisk.*

Bevis. Detta gäller eftersom $(A^{-1})^T = (A^T)^{-1} = A^{-1}$. $\diamondsuit$

LDU-faktoriseringen av en symmetrisk, inverterbar matris är speciellt enkel:

Sats 3.12. *Om A är symmetrisk och inverterbar och kan LDU-faktoriseras, så har faktoriseringen utseendet $A = LDL^T$, dvs. $U = L^T$.*

Bevis. Genom transponering av $A = LDU$ fås att

$$A = A^T = (LDU)^T = U^T D^T L^T = U^T D L^T ,$$

vilket är en LDU-faktorisering av A. Entydigheten hos LDU-faktoriseringar (Sats 3.8) ger att $U^T = L$ och $L^T = U$. $\diamondsuit$

Avrundningsfel. Partiell pivotering

Då man löser linjära ekvationssystem, utförs ett stort antal räkneoperationer. För ett system med n ekvationer och n obekanta krävs ungefär $2n^3/3$ räkneoperationer för att få fram lösningen. Vid varje räkneoperation måste vi räkna med ett avrundningsfel (koefficienterna avrundas kanske till 3 signifikanta siffror) och detta påverkar naturligtvis lösningen, som då bara blir approximativ.

Ibland kan koefficientmatrisen ha en inneboende *känslighet*, som gör att en liten ändring av ett matriselement i räkneschemat, ett avrundningsfel, **oundvikligen** får stor inverkan på (den approximativa) lösningen. En annan felkälla kan vara att kalkylerna görs på ett sådant sätt att felen förstoras i onödan. Vi skall om en stund beskriva en enkel metod att undvika sådana fel, som inte beror på genuin känslighet.

Följande är ett exempel på en koefficientmatris som är genuint känslig:

Exempel 3.14. Av de två systemen

$$\begin{cases} u + v = 2 \\ u + 1,0001v = 2 \end{cases} \quad \text{och} \quad \begin{cases} u + v = 2 \\ u + 1,0001v = 2,0001 \end{cases}$$

med samma koefficientmatris, har det första systemet lösningen $u = 2$, $v = 0$, medan det andra har lösningen $u = 1$, $v = 1$. Orsaken är helt enkelt att de två räta linjerna $u + v = \alpha$ och $u + 1,0001v = \beta$ är **nästan parallella**, varför en ändring i femte siffran i ett tal i högerledet förorsaker en ändring i första siffran i skärningspunkten (dvs. lösningen). Eliminering ger

$$\begin{pmatrix} 1 & 1 & | & \alpha \\ 1 & 1,0001 & | & \beta \end{pmatrix} \to \begin{pmatrix} 1 & 1 & | & \alpha \\ 0 & 0,0001 & | & \beta - \alpha \end{pmatrix}.$$

"Parallelliteten" hos de två linjerna visar sig här i att det uppstår ett pivotelement, som är nära noll. I fortsättningen av kalkylerna kommer vi att dividera med detta tal nära noll, och därför förstoras eventuella avrundningsfel. Här hjälper det inte heller att permutera ekvationerna och därför är känsligheten genuin.

Om det är möjligt, bör man undvika att göra kalkylerna så att något pivotelement blir ett tal nära noll. Metoden med *partiell pivotering* anger en enkel regel för detta:

Varje gång man skall eliminera med hjälp av $(BO1^+)$ för att få ett system i echelonform, bör man först permutera de återstående raderna så, att det till beloppet största elementet i resten av kolonnen blir pivotelement.

I koefficientmatrisen (= vänstra delen av ett räkneschema)

$$\begin{pmatrix} d_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ 0 & a_{32} & \cdots & * \\ - & - & - & - \\ 0 & a_{m2} & \cdots & * \end{pmatrix}$$

har vi redan ett pivotelement, d_{11} $(\neq 0)$. I nästa steg bör vi alltså bland talen a_{22}, a_{32}, $\ldots$, a_{m2} välja det som är störst till beloppet och sedan låta den rad, som talet ingår i, bli rad nummer två så att det utvalda talet blir nästa pivotelement (ifall inte alla dessa tal är noll, varvid man naturligtvis går vidare till tredje kolonnen).

Övningsuppgifter

1. Vad händer med en 3/3-matris A om man multiplicerar den (a) från vänster, (b) från höger med matriserna

$$E_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \quad \text{respektive} \quad E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}?$$

2. I ett ekvationssystem med tre ekvationer och tre obekanta gausseliminerades genom att i tur och ordning basoperationerna

$$\begin{array}{rl} (i) & (2) \to (2) - 4\,(1) \\ (ii) & (3) \to (3) - 2\,(1) \\ (iii) & (3) \to (3) - \dfrac{1}{3}(2) \end{array}$$

användes, varvid det triangulära systemet

$$\begin{aligned} u - 2v + \ 3w &= \ 11 \\ 9v - 13w &= -40 \\ \frac{4}{3}w &= \ \frac{4}{3} \end{aligned}$$

uppstod. Hur såg systemet ut från början?

3. Lös ekvationen $A\mathbf{x} = \mathbf{b}$, då $A = LU$ och

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}.$$

4. Antag att P_{ij} och P_{kl} är enkla permutationsmatriser av samma typ. När kommuterar de, dvs. när är $P_{ij}P_{kl} = P_{kl}P_{ij}$?

5. Hur många permutationsmatriser finns det (a) av typ $3/3$, (b) av typ n/n?

6. Finn faktorerna L, och U i LU-faktoriseringen av matrisen

$$(a) \quad A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \qquad (b) \quad A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \\ -1 & 2 & 1 \end{pmatrix}.$$

7. Vilka värden på a och b leder till radbyte vid triangulering av A och vilka värden gör A singulär, då

$$A = \begin{pmatrix} 1 & 2 & 2 \\ a & 8 & 3 \\ 0 & b & 3 \end{pmatrix}.$$

8. Bestäm LDU-faktoriseringen av PA, där P är en lämplig permutationsmatris, då

$$(a) \quad A = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}, \qquad (b) \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 2 \\ 1 & 1 & 1 \end{pmatrix}, \qquad (c) \quad A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & -1 \\ 3 & 2 & 0 \end{pmatrix}.$$

9. Bestäm en permutationsmatris P sådan att PA har en LDU-faktorisering och beräkna denna, då

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ -2 & -4 & 5 & -2 \\ 2 & 6 & 4 & -1 \\ -1 & 0 & 4 & -3 \end{pmatrix}.$$

10. Antag att LU-faktoriseringen $A = LU$ är given men att systemet som skall lösas är $A^T\mathbf{y} = \mathbf{b}$. Finn en ändamålsenlig metod att beräkna $\mathbf{y}$.

11. Bestäm inversen till matrisprodukten

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

12. Beräkna inverserna, ifall de existerar, till

$$(a) \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 1 & 1 \end{pmatrix}, \qquad (b) \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 5 & 9 & 1 & 6 \end{pmatrix}.$$

13. Antag att A, B och C är inverterbara matriser. Vad är inversen till $AB^{-1}C$? Är A^2 inverterbar? Är $A + B$ inverterbar? Visa att $A^{-1}(A + B)B^{-1} = A^{-1} + B^{-1}$.

14. Bestäm inversen till

$$(a) \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad (b) \quad \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

15. Finn alla 2/2-matriser, som är sin egen invers.

16. Ange de värden på a för vilka nedanstående matris saknar invers:

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

17. Visa att det finns oändligt många *vänsterinverser* till matrisen

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

(dvs. matriser B sådana att $BA = I$) men ingen *högerinvers* till A (dvs. ingen matris C sådan att $AC = I$).

18. Bestäm den symmetriska LDU-faktoriseringen $A = LDL^T$ av

$$(a) \quad A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 6 & 4 \\ 0 & 4 & 11 \end{pmatrix}, \qquad (b) \quad A = \begin{pmatrix} 1 & 3 & 5 \\ 3 & 12 & 18 \\ 5 & 18 & 30 \end{pmatrix}.$$

19. Visa att produkten av två uppåt triangulära n/n-matriser med ettor i diagonalen är en matris av samma typ.

20. En kvadratisk, inverterbar matris sägs vara *ortogonal* om $A^{-1} = A^T$. Visa, (a) att produkten av två ortogonala matriser är en ortogonal matris, (b) att om A är ortogonal så är också A^{-1} ortogonal.

21. Antag att A är en kvadratisk, icke-singulär matris. Visa att både A och A^T ger upphov till samma pivotelement, om enbart $(BO1^+)$ används.

22. Antag att A är en inverterbar matris och att matriserna A och B kommuterar, dvs. att $AB = BA$. Visa

    (a) att också A^{-1} och B kommuterar,

    (b) att om A och B dessutom är symmetriska, så är också $A^{-1}B$ symmetrisk.

23. Visa att varje idempotent matris A, som inte är I, måste vara singulär (A är idempotent om $A^2 = A$).

24. Räkna ut LU-faktoriseringen av matrisen

$$(a) \quad A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 1 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}, \qquad (b) \quad A = \begin{pmatrix} 1 & 2 & 0 & 2 & 1 \\ -1 & -2 & 1 & 1 & 0 \\ 1 & 2 & -3 & -7 & -2 \end{pmatrix}.$$

25. Bestäm matrisen A då vi vet att A är inverterbar och att inversen till matrisen

$$\begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix} A \qquad \text{är} \qquad 7 \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}.$$

26. Låt A vara en n/n-matris, som har talen $d_1, \dots, d_n$ i diagonalen men i övrigt består ettor. Antag att $d_i > 1$ för varje i. Visa att $x^T A x > 0$ för varje $n/1$-vektor $x \neq 0$ och slut härav att A är icke-singulär. (Ledning: Skriv A som summan av en diagonalmatris med talen $d_i - 1$ i diagonalen och en matris som består av enbart ettor.)

27. En *kubisk ri-funktion* $f(x)$ genom givna punkter $(x_0, y_0), \dots, (x_n, y_n)$ är definierad i varje intervall $[x_{i-1}, x_i]$ $(i = 1, \dots, n)$ som ett tredjegradspolynom $q_i(x)$. Dessa polynom är valda så att $f(x)$, $f'(x)$ och $f''(x)$ blir kontinuerliga i skarvningspunkterna $x_1, \dots, x_{n-1}$. Om man dessutom kräver att $f''(x_0) = f''(x_n) = 0$ så sägs den kubiska ri-funktionen vara *naturlig*. Om alla intervall har samma längd $h = x_i - x_{i-1}$, $i = 1, \dots, n$, så är för den naturliga kubiska ri-funktionen

$$q_i(x) = ty_i + (1-t)y_{i-1} + ht(1-t)\left[(k_{i-1} - d_i)(1-t) - (k_i - d_i)t\right],$$

där $x = x_{i-1} + th$, $d_i = (y_i - y_{i-1})/h$ och talen $k_0, \dots, k_n$ fås ur matrisekvationen

$$\begin{pmatrix} 2 & 1 & 0 & .. & 0 & 0 \\ 1 & 4 & 1 & .. & 0 & 0 \\ 0 & 1 & 4 & .. & 0 & 0 \\ - & - & - & - & - & - \\ 0 & 0 & 0 & .. & 4 & 1 \\ 0 & 0 & 0 & .. & 1 & 2 \end{pmatrix} \begin{pmatrix} k_0 \\ k_1 \\ k_2 \\ - \\ k_{n-1} \\ k_n \end{pmatrix} = 3 \begin{pmatrix} d_1 \\ d_1 + d_2 \\ d_2 + d_3 \\ - \\ d_{n-1} + d_n \\ d_n \end{pmatrix}.$$

(a) Bestäm LU-faktoriseringen av koefficientmatrisen i det fall att den är en 4/4-matris. (b) Bestäm den naturliga kubiska ri-funktionen genom $(0,1)$, $(1,3)$ och $(2,2)$.

4. Vektorrum

Tidigare har vi räknat upp en rad av räkneregler som gäller för m/n-matriser. Dessa regler gäller inte bara för varje matristyp m/n utan också för många andra "objekt" som t.ex. funktioner, talföljder, matrisföljder osv. För att inte behöva upprepa alla bevis för varje typ av objekt, inför vi begreppet vektorrum som en mängd av abstrakta objekt (vektorer), som kan representera vilket som helst av de ovannämnda slagen av objekt. Allt som kan bevisas för vektorer i vektorrum, kommer då automatiskt att gälla för varje slag av objekt som underlyder vissa gemensamma grundräkneregler som vi kallar axiom.

På samma sätt har ju det abstrakta talet 2 införts för att man inte skall behöva ha skilda bevis för multiplikation med två liter och för multiplikation med två grodor.

Axiom för vektorrum

Det matematiska begreppet vektorrum är centralt i den linjära algebran. Vi inför det genom att räkna upp en rad av axiom (= räkneregler) som ett vektorrum skall uppfylla. Dessa axiom är ett urval bland de räkneregler som vi tidigare har konstaterat att gäller för matriser.

Definition 4.1. En mängd E sägs vara ett *vektorrum* och elementen i E kallas *vektorer*, om två räkneoperationer, *addition* och *multiplikation med skalär*, är definierade i E så att de nedan uppräknade axiomen gäller. Additionen tillordnar varje par $\mathbf{x}$, $\mathbf{y}$ i E en summa $\mathbf{x}+\mathbf{y}$ i E och multiplikationen med skalär tillordnar varje par λ, $\mathbf{x}$ ($\lambda \in \mathbf{R}$, $\mathbf{x} \in E$) en produkt $\lambda\mathbf{x}$ i E, så att följande gäller:

I. (*a*) $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x} \quad (\mathbf{x}, \mathbf{y} \in E)$, (kommutationslag)

(*b*) $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z} \quad (\mathbf{x}, \mathbf{y}, \mathbf{z} \in E)$, (associationslag)

(*c*) Det finns ett element i $\mathbf{0} \in E$, kallat *nollvektorn*, sådant att $\mathbf{x}+\mathbf{0}=\mathbf{x}$

(*d*) För varje $\mathbf{x} \in E$ finns ett element $-\mathbf{x} \in E$, den *motsatta vektorn*, sådant att $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$

II. (*a*) $\lambda(\mu\mathbf{x})=(\lambda\mu)\mathbf{x}, \quad (\lambda, \mu \in \mathbf{R}, \mathbf{x} \in E)$, (associationslag)

(*b*) $1\mathbf{x}=\mathbf{x} \quad (\mathbf{x} \in E)$

III. (*a*) $(\lambda+\mu)\mathbf{x}=\lambda\mathbf{x}+\mu\mathbf{x}, \quad (\lambda, \mu \in \mathbf{R}, \mathbf{x} \in E)$, (distributionslag)

(*b*) $\lambda(\mathbf{x}+\mathbf{y})=\lambda\mathbf{x}+\lambda\mathbf{y}, \quad (\lambda \in \mathbf{R}, \mathbf{x}, \mathbf{y} \in E)$, (distributionslag).

Vi ger genast några exempel på mängder som är vektorrum:

Exempel 4.1. Mängden E av alla m/n-matriser, försedd med vanlig matrisaddition och vanlig multiplikation av en matris med en skalär, är ett vektorrum, eftersom de ovan uppräknade axiomen lätt kan konstateras gälla för dessa operationer. Speciella exempel på vektorrum är följaktligen mängden $\mathbf{R}^m$ av alla $m/1$-matriser, dvs. kolonnvektorer, och mängden $\mathbf{R}^n$ av alla $1/n$-matriser, dvs. radvektorer. Man brukar använda samma beteckning för bägge, dvs. $\mathbf{R}^m$ och $\mathbf{R}^n$, på grund av att skillnaden mellan en kolonnvektor och en radvektor är obetydlig: Den ena typen övergår ju i den andra genom transponering (det är bara vid matrismultiplikation som det är väsentligt hur man skriver en vektor).

Då m och n är 2 eller 3 representerar vektorrummen $\mathbf{R}^2$ och $\mathbf{R}^3$ på känt sätt ett två- respektive ett tredimensionellt "koordinatsystem".

Exempel 4.2. Mängden $\mathcal{M}$ av alla funktioner $f : \mathbf{R} \to \mathbf{R}$ med definitionsmängden $[a, b]$ är ett vektorrum då följande räkneoperationer är definierade i $\mathcal{M}$:

I. Om $f_1 \in \mathcal{M}$ och $f_2 \in \mathcal{M}$ så definieras $f_1 + f_2$ genom
$$(f_1 + f_2)(x) = f_1(x) + f_2(x)\,, \quad x \in [a, b]\,;$$

II. Om $f \in \mathcal{M}$ och $\lambda \in \mathbf{R}$, definieras λf genom
$$(\lambda f)(x) = \lambda f(x)\,, \quad x \in [a, b]\,.$$

Både $f_1 + f_2$ och λf är funktioner i mängden $\mathcal{M}$. Att alla axiomen är uppfyllda är lätt att verifiera (på grund av att de reella talen uppfyller axiomen). T.ex. är nollvektorn i axiom I (c) precis den funktion (betecknad med 0 och kallad *nollfunktionen*), som har värdet 0 i hela intervallet $[a, b]$. Att axiom I (d) är uppfyllt ser vi av följande: Om $f \in \mathcal{M}$ så väljer vi $-f$ att vara $(-1)f$. Då är $f + (-f) = 0$, ty

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0\,, \quad x \in [a, b]\,.$$

Exempel 4.3. Mängden E av alla funktioner $f : \mathbf{R} \to \mathbf{R}^n$, med definitionsmängden $[a, b]$ och med värden i vektorrummet $\mathbf{R}^n$, är ett vektorrum då operationerna definieras som i föregående exempel. Axiomen är lätta att verifiera på grund av att värdemängden $\mathbf{R}^n$ är ett vektorrum.

Utgående från axiomen kan hela teorin byggas upp systematiskt. Vi bevisar här som exempel några grundläggande saker:

1. *Nollvektorn är entydig.*

Om både $\mathbf{0}_1$ och $\mathbf{0}_2$ är nollvektorer i ett vektorrum, så gäller:

$$\begin{aligned} \mathbf{x} + \mathbf{0}_1 = \mathbf{x} \text{ för varje } \mathbf{x} \quad &\Longrightarrow \quad \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2 \\ \mathbf{x} + \mathbf{0}_2 = \mathbf{x} \text{ för varje } \mathbf{x} \quad &\Longrightarrow \quad \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1\,. \end{aligned}$$

Eftersom additionen är kommutativ, är därmed $\mathbf{0}_1 = \mathbf{0}_2$.

2. *Vektorn* $-\mathbf{x}$ *är entydig, då vektorn* $\mathbf{x}$ *är given.*

   Om både $(-\mathbf{x})_1$ och $(-\mathbf{x})_2$ är motsatta vektorer till $\mathbf{x}$, är $(-\mathbf{x})_1+\mathbf{x} = \mathbf{0}$ och $\mathbf{x}+(-\mathbf{x})_2 = \mathbf{0}$ och därmed är de lika:

$$(-\mathbf{x})_1 = (-\mathbf{x})_1 + (\mathbf{x} + (-\mathbf{x})_2) = ((-\mathbf{x})_1 + \mathbf{x}) + (-\mathbf{x})_2 = (-\mathbf{x})_2 \,.$$

   På grund av denna entydighet och kommutativiteten kommer $\mathbf{x}$ och $-\mathbf{x}$ att vara varandras motsatta vektorer, så att

$$-(-\mathbf{x}) = \mathbf{x} \,.$$

3. *För givna* $\mathbf{x}$ *och* $\mathbf{y}$ *har ekvationen* $\mathbf{x} + \mathbf{t} = \mathbf{y}$ *den entydiga lösningen* $\mathbf{t} = \mathbf{y} + (-\mathbf{x})$.

   På grund av associativiteten behöver man inte skriva ut parenteser i summor. Vi får då följande ekvivalenser genom att addera $-\mathbf{x}$ (eller $\mathbf{x}$) till bägge leden:

$$\mathbf{x} + \mathbf{t} = \mathbf{y} \quad \Leftrightarrow \quad (-\mathbf{x}) + \mathbf{x} + \mathbf{t} = (-\mathbf{x}) + \mathbf{y} \quad \Leftrightarrow \quad \mathbf{t} = (-\mathbf{x}) + \mathbf{y} \,.$$

   I fortsättningen kommer vi oftast att använda beteckningen $\mathbf{y} - \mathbf{x}$ i stället för $\mathbf{y} + (-\mathbf{x})$.

4. *Det gäller att* $\lambda\mathbf{x} = \mathbf{0}$ *om och endast om* $\lambda = 0$ *eller* $\mathbf{x} = \mathbf{0}$.

   Först ses att $0\mathbf{x} = \mathbf{0}$, ty om man löser ut $0\mathbf{x}$ ur

$$\mathbf{x} + 0\mathbf{x} = (1 + 0)\mathbf{x} = 1\mathbf{x} = \mathbf{x} \,,$$

   får man att $0\mathbf{x} = \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ (enligt 3). Med hjälp av detta fås $\mathbf{x} + (-1)\mathbf{x} = (1 + (-1))\mathbf{x} = 0\mathbf{x} = \mathbf{0}$, varför

$$(-1)\mathbf{x} = -\mathbf{x}$$

   enligt 2. Detta ger att

$$\lambda\mathbf{0} = \lambda(\mathbf{x} + (-1)\mathbf{x})) = (\lambda\mathbf{x}) + (-\lambda\mathbf{x}) = \mathbf{0} \,.$$

   Antag omvänt att $\lambda\mathbf{x} = \mathbf{0}$. Om nu $\lambda \neq 0$, så är $\mathbf{x} = (1/\lambda)(\lambda\mathbf{x}) = (1/\lambda)\mathbf{0} = \mathbf{0}$. Är igen $\mathbf{x} \neq \mathbf{0}$, så är $\lambda = 0$ (vore nämligen $\lambda \neq 0$ så vore $\mathbf{x} = \mathbf{0}$ enligt föregående mening). Minst en av faktorerna måste alltså vara noll om $\lambda\mathbf{x} = \mathbf{0}$.

Underrum

Ofta ligger ett vektorrum inne i ett annat vektorrum. Om bägge vektorrummen har samma addition och multiplikation med skalär, säger man att det "mindre" vektorrummet är ett underrum av det "större". Nedanstående definition är mera praktisk att använda men innebär ändå precis detta.

Definition 4.2. Antag att E är ett vektorrum. En delmängd U av E är ett *underrum* av E om mängden U är *sluten* med avseende på operationerna addition och multiplikation med skalär i följande mening:

$$\mathbf{x}, \mathbf{y} \in U;\ \lambda \in \mathbf{R} \quad \Longrightarrow \quad \left\{ \begin{array}{c} \mathbf{x} + \mathbf{y} \in U \\ \lambda\mathbf{x} \in U \end{array} \right\} .$$

Eftersom axiomen för vektorrum gäller i E, så gäller de automatiskt också i underrummet U. Således gäller: **Varje underrum av E är ett vektorrum**. Ett underrum av ett vektorrum E är alltså ett annat vektorrum som råkar finnas inne i E.

Exempel 4.4. Betrakta i vektorrummet $\mathbf{R}^3$ delmängden U av alla vektorer av formen $(\, x_1 \quad 0 \quad x_3 \,)$, dvs. x_1x_3-planet. Då är mängden U sluten med avseende på räkneoperationerna, ty summan av två vektorer, vilkas andra komponent är 0, har andra komponenten 0, och produkten av en skalär λ med en vektor, vilkens andra komponent är 0, har 0 som andra komponent. Alltså är x_1x_3-planet ett underrum av $\mathbf{R}^3$.

Exempel 4.5. Sätt $U = \{(\, x \quad y \,) \in \mathbf{R}^2 \mid x + y = 0\}$ och förse U med samma räkneoperationer som i $\mathbf{R}^2$ (dvs. addition och multiplikation med skalär). Geometriskt är U en rät linje genom origo. Om man får som uppgift att visa att U är ett vektorrum, räcker det att visa att U är ett underrum av $\mathbf{R}^2$, som vi redan vet att är ett vektorrum. Vi undersöker därför om U är slutet med avseende på räkneoperationerna:

Tag $\mathbf{u} = (\, x_1 \quad y_1 \,) \in U$ och $\mathbf{v} = (\, x_2 \quad y_2 \,) \in U$. Då är $x_1 + y_1 = 0$ och $x_2 + y_2 = 0$ och summan av komponenterna i vektorerna

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (\, x_1 + x_2 \quad y_1 + y_2 \,) \\ \lambda\mathbf{u} &= (\, \lambda x_1 \quad \lambda y_1 \,)\,, \quad (\lambda \in \mathbf{R}), \end{aligned}$$

blir då noll, varför de enligt definitionen är vektorer i U:

$$\begin{aligned} (x_1 + x_2) + (y_1 + y_2) &= (x_1 + y_1) + (x_2 + y_2) = 0\,, \\ (\lambda x_1) + (\lambda y_1) &= \lambda(x_1 + y_1) = 0\,. \end{aligned}$$

Följaktligen är U ett underrum av $\mathbf{R}^2$ och därmed också ett vektorrum.

Exempel 4.6. Om vi sätter $U_1 = \{(\, x \quad y \,) \in \mathbf{R}^2 \mid x + y = 1\}$ så är U_1 inget underrum av $\mathbf{R}^2$. Mängden U_1 är nämligen inte sluten med avseende på t.ex. multiplikation med skalär: Om $(\, x \quad y \,) \in U_1$, så är $x + y = 1$ och då är $2\,(\, x \quad y \,) \notin U_1$, eftersom $(2x) + (2y) = 2(x + y) = 2$.

Allmänt gäller att ett underrum måste innehålla nollvektorn $\mathbf{0}$, eftersom varje underrum ju är ett vektorrum.

Om E är ett vektorrum, så är delmängderna $\{\mathbf{0}\}$ och E underrum av E. Dessa kallas *triviala* underrum. *Äkta* underrum kallas varje underrum, som inte omfattar hela vektorrummet E.

Det är lätt att inse att en delmängd U av $\mathbf{R}^2$ är ett icke-trivialt underrum om och endast om U är en **rät linje genom origo**. På samma sätt är en delmängd V av $\mathbf{R}^3$ är ett icke-trivialt underrum, om och endast om V är en **rät linje genom origo** eller är ett **plan genom origo**.

Exempel 4.7. Mängden $\mathcal{P}$ av alla polynomfunktioner är ett underrum av vektorrummet $\mathcal{M}$ (i exempel 4.2) av alla funktioner, som är definierade på intervallet $[a,\, b]$. Mängden $\mathcal{P}_n$ av alla polynomfunktioner av högst graden n är ett underrum av både $\mathcal{P}$ och $\mathcal{M}$.

Kolonnrummet och nollrummet till en matris

Med varje matris kommer vi att associera fyra underrum. I detta avsnitt introducerar vi två av dem.

Redan tidigare har vi sett, att i en matrisekvation $A\mathbf{x} = \mathbf{b}$ kan vänsterledet $A\mathbf{x}$ uppfattas som en linjärkombination av kolonnvektorerna i matrisen A. I t.ex. i ekvationen

$$\begin{pmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

kan vänsterledet skrivas

$$u \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} + v \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix} \qquad \text{(Se Sats 2.1)}\,.$$

Den generella definitionen är:

Definition 4.3. Antag att $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ är vektorer i ett vektorrum E (dvs. element i mängden E). En *linjärkombination* av $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ är en vektor av formen

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p\,,$$

där koefficienterna $c_1, c_2, \ldots, c_p$ är reella tal. Genom att variera dessa koefficienter fås **alla** linjärkombinationer av vektorerna $\mathbf{v}_i$, $i = 1, \ldots, p$. Mängden av alla dessa,

$$U = \operatorname{spn}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} = \{c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p \mid c_1, \ldots, c_p \in \mathbf{R}\}\,,$$

kallas *spannet* av $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$. Man säger att spannet U *spänns upp* av vektorerna $\mathbf{v}_i$, $i = 1, \ldots, p$.

Sats 4.1. *Om $\mathbf{v}_1, \ldots, \mathbf{v}_p$ är vektorer i ett vektorrum E, så är spannet*

$$U = \operatorname{spn}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$$

ett underrum av E.

Bevis. Om $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ och $\mathbf{y} = d_1\mathbf{v}_1 + \cdots + d_p\mathbf{v}_p$ är godtyckliga vektorer i U och om $\lambda \in \mathbf{R}$, så är

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (c_1 + d_1)\mathbf{v}_1 + \cdots + (c_p + d_p)\mathbf{v}_p \in U\,, \\ \lambda\mathbf{x} &= (\lambda c_1)\mathbf{v}_1 + \cdots + (\lambda c_p)\mathbf{v}_p \in U\,. \end{aligned}$$

Enligt definitionen är U då ett underrum av E. $\diamondsuit$

Antag nu att $A = (\,\mathbf{a}_1 \ldots \mathbf{a}_n\,)$ är en m/n-matris med kolonnvektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Definition 4.4. *Kolonnrummet* $R(A)$ till A är spannet av kolonnerna i A, dvs.

$$R(A) = \operatorname{spn}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}\,.$$

Exempel 4.8. Kolonnrummet $R(A)$ till matrisen

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

är spannet av kolonnerna $(\,2 \quad 1 \quad 0\,)^T$ och $(\,0 \quad 1 \quad 2\,)^T$. Geometriskt är $R(A)$ det plan i $\mathbf{R}^3$ som går genom origo och som innehåller dessa två vektorer.

Ibland behövs inte alla kolonner i matrisen för att beskriva kolonnrummet. T.ex. i matrisen

$$B = (\,\mathbf{b}_1 \quad \mathbf{b}_2\,) = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$$

är $\mathbf{b}_1 = 2\mathbf{b}_2$. En godtycklig linjärkombination av $\mathbf{b}_1$ och $\mathbf{b}_2$ kan därför skrivas som en linjärkombination av enbart $\mathbf{b}_2$:

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = (2c_1 + c_2)\mathbf{b}_2\,.$$

Alltså är $R(B) = \operatorname{spn}\{\mathbf{b}_1, \mathbf{b}_2\} = \operatorname{spn}\{\mathbf{b}_2\}$.

Giltigheten hos det första påståendet i nästa sats följer ur Sats 4.1 medan det andra är en följd av att vänstra ledet i ekvationen $A\mathbf{x} = \mathbf{b}$ är en linjärkombination av A:s kolonner:

Sats 4.2. *Antag att A är en m/n-matris. Då gäller:*

(*i*) *$R(A)$ är ett underrum av $\mathbf{R}^m$;*

(*ii*) *Systemet $A\mathbf{x} = \mathbf{b}$ har en lösning* (dvs. är konsistent) *om och endast om $\mathbf{b} \in R(A)$.*

Vi har sett att ett homogent system $A\mathbf{x} = \mathbf{0}$ alltid har åtminstone den triviala lösningen $\mathbf{x} = \mathbf{0}$. Om det existerar fria variabler så finns det dessutom andra lösningar.

Definition 4.5. Låt A vara en m/n-matris. Mängden av alla lösningar till den homogena ekvationen $A\mathbf{x} = \mathbf{0}$, dvs.

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

kallas *nollrummet* till matrisen A.

Sats 4.3. *Nollrummet $N(A)$ till en m/n-matris är ett underrum av $\mathbf{R}^n$.*

Bevis. Antag att $\mathbf{x}$ och $\mathbf{y}$ är vektorer i $N(A)$ och låt $\lambda \in \mathbf{R}$ vara godtyckligt. Då är $A\mathbf{x} = \mathbf{0}$ och $A\mathbf{y} = \mathbf{0}$, och därmed är

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= A\mathbf{x} + A\mathbf{y} = \mathbf{0} \quad \text{och} \\ A(\lambda\mathbf{x}) &= \lambda A\mathbf{x} = \mathbf{0}\,. \end{aligned}$$

Alltså är $\mathbf{x} + \mathbf{y} \in N(A)$ och $\lambda\mathbf{x} \in N(A)$, varför $N(A)$ är ett underrum (av $\mathbf{R}^n$). $\diamondsuit$

Exempel 4.9. Nollrummet till en matris

$$A = \begin{pmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & 0 & 1 \\ 0 & -5 & 6 & 1 \end{pmatrix}$$

kan framställas i form av ett spann genom att man löser ekvationen $A\mathbf{x} = \mathbf{0}$. Först överförs räkneschemat i reducerad echelonform:

$$A \to \cdots \to \begin{pmatrix} 1 & 0 & 3/5 & 3/5 \\ 0 & 1 & -6/5 & -1/4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

De fria variablerna ges godtyckliga värden, $x_3 = s$, $x_4 = t$, och lösningen skrivs ut slutgiltigt i vektorform:

$$\begin{array}{rl} x_1 = & -\frac{3}{5}s - \frac{3}{5}t \\ x_2 = & \frac{6}{5}s + \frac{1}{5}t \\ x_3 = & s \\ x_4 = & \qquad t \end{array}, \qquad \text{dvs.} \quad \mathbf{x} = \frac{s}{5}\begin{pmatrix} -3 \\ 6 \\ 5 \\ 0 \end{pmatrix} + \frac{t}{5}\begin{pmatrix} -3 \\ 1 \\ 0 \\ 5 \end{pmatrix}.$$

En vektor $\mathbf{x}$ är alltså i $N(A)$ om och endast om den är en linjärkombination av vektorerna $\mathbf{v}_1 = (\,-3 \quad 6 \quad 5 \quad 0\,)^T$ och $\mathbf{v}_2 = (\,-3 \quad 1 \quad 0 \quad 5\,)^T$. Alltså är

$$N(A) = \operatorname{spn}\{\mathbf{v}_1, \mathbf{v}_2\}.$$

Strukturen hos lösningen till ett system

Ett exempel visar den allmänna strukturen hos mängden av lösningar till ett system:

Exempel 4.10. Betrakta systemet $A\mathbf{x} = \mathbf{b}$, där

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 5 \\ 5 \end{pmatrix}.$$

Vi löser systemet med hjälp av basoperationer och finner den reducerade echelonformen,

$$(A \mid \mathbf{b}) \to \cdots \to \left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & -2 \\ 0 & 0 & 1 & 1/3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Variablerna x_2 och x_4 är fria och ges godtyckliga värden, $x_2 = s$ och $x_4 = t$, och vi avläser att

$$\begin{array}{rl} x_1 = & -2 - 3s - t \\ x_2 = & \qquad s \\ x_3 = & 1 \qquad -\frac{1}{3}t \\ x_4 = & \qquad\qquad t. \end{array}$$

I vektorform blir detta

$$\mathbf{x} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{t}{3} \begin{pmatrix} -3 \\ 0 \\ -1 \\ 3 \end{pmatrix}.$$

Om vi skriver lösningen $\mathbf{x}$ i den mer hanterliga formen $\mathbf{x} = \mathbf{x}_0 + s\mathbf{y}_1 + t\mathbf{y}_2$, så representerar linjärkombinationen $s\mathbf{y}_1 + t\mathbf{y}_2$ vilket som helst element i $N(A)$, dvs. en godtycklig lösning (den allmänna lösningen) till den homogena ekvationen $A\mathbf{x} = \mathbf{0}$ (om $\mathbf{b} = \mathbf{0}$ så blir ju $\mathbf{x}_0 = \mathbf{0}$). Den första termen $\mathbf{x}_0$ är en speciell lösning (en partikulärlösning) till $A\mathbf{x} = \mathbf{b}$ som fås då man sätter $s = 0$ och $t = 0$. Mängden $\mathbf{x}_0 + N(A)$ av alla lösningar är alltså underrummet $N(A)$ förskjutet bort från origo O med hjälp av en vektor $\mathbf{x}_0$:

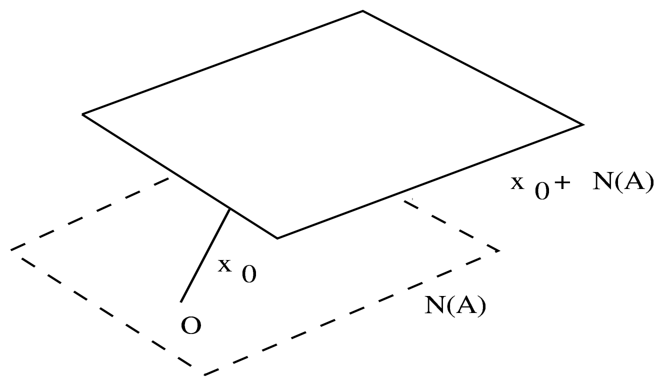

fig. 1

Denna struktur återkommer hos mängden an lösningar till varje ekvationssystem men låt oss först slå fast definitionerna:

Definition 4.6. Med den *allmänna lösningen* till en ekvation $A\mathbf{x} = \mathbf{b}$ avses ett lösningsuttryck innehållande parametrar $s, t, \ldots$, sådant att **alla** lösningar till ekvationen kan fås genom att man varierar dessa parametrars värden. En *partikulärlösning* är någon speciell lösning till $A\mathbf{x} = \mathbf{b}$.

Sats 4.4. *Den allmänna lösningen till en konsistent ekvationen* $A\mathbf{x} = \mathbf{b}$ *har formen* $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$, *där* $\mathbf{x}_0$ *är någon* (vilken som helst) *partikulärlösning till* $A\mathbf{x} = \mathbf{b}$ *och* $\mathbf{y}$ *är den allmänna lösningen till motsvarande homogena ekvation* $A\mathbf{x} = \mathbf{0}$.

Bevis. Det framgår av lösningsproceduren (se exempel 4.10) att varje lösning har den form som satsen anger. För att vi skall se att partikulärlösningen $\mathbf{x}_0$ kan vara vilken lösning som helst till $A\mathbf{x} = \mathbf{b}$, byter vi ut $\mathbf{x}_0$ mot en annan partikulärlösning $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}_1$ ($\mathbf{y}_1 \in N(A)$) genom att skriva uttrycket för $\mathbf{x}$ i formen $\mathbf{x} = \mathbf{x}_1 + (\mathbf{y} - \mathbf{y}_1)$, där $\mathbf{y} - \mathbf{y}_1$ nu är en lösning till $A\mathbf{x} = \mathbf{0}$. $\diamondsuit$

Linjärt beroende och oberoende

Exempel 4.11. Betrakta t.ex. vektorerna

$$\mathbf{a}_1 = (\,1 \quad -1 \quad 2\,)\,, \quad \mathbf{a}_2 = (\,1 \quad 1 \quad 0\,)\,, \quad \mathbf{a}_3 = (\,2 \quad 0 \quad 2\,)$$

i $\mathbf{R}^3$. Mellan dessa gäller ett linjärt samband i den meningen att $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$, vilket också kan skrivas $\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_3 = \mathbf{0}$. Dessa vektorer sägs därför vara linjärt beroende (av varandra).

Definition 4.7. Vektorer $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ i ett vektorrum E (eller mängden $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}$ av dessa vektorer) är *linjärt beroende* om det existerar ett linjärt samband

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_p\mathbf{a}_p = \mathbf{0}\,, \tag{1}$$

där minst en av koefficienterna c_i är **olik noll.** Vektorerna är *linjärt oberoende* om de inte är linjärt beroende. Detta innebär att (1) inte kan gälla med mindre än att alla koefficienterna är noll:

$$\begin{cases} \qquad \mathbf{a}_1, \ldots, \mathbf{a}_p \text{ är linjärt oberoende om och endast om} \\[2ex] c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_p\mathbf{a}_p = \mathbf{0} \quad \Longrightarrow \quad c_1 = c_2 = \cdots = c_p = 0\,. \end{cases} \tag{2}$$

Om vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_p$ är linjärt beroende så gäller (1), där minst en av koefficienterna är olik noll. Om t.ex. $c_p \neq 0$ så är

$$\mathbf{a}_p = -\frac{c_1}{c_p}\mathbf{a}_1 - \cdots - \frac{c_{p-1}}{c_p}\mathbf{a}_{p-1}\,,$$

dvs. en linjärkombination av de andra vektorerna. Detta gäller generellt: **Om vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_p$ är linjärt beroende så gäller om $p > 1$, att minst en av vektorerna är en linjärkombination av de andra, och om $p = 1$, att $\mathbf{a}_1 = \mathbf{0}$.**

Exempel 4.12. Då vi skall undersöka om ett antal vektorer

$$\mathbf{a}_1 = (\,1 \quad 1 \quad 1\,)\,, \quad \mathbf{a}_2 = (\,1 \quad 0 \quad 1\,)\,, \quad \mathbf{a}_3 = (\,0 \quad 2 \quad 1\,)\,,$$

är linjärt beroende eller oberoende använder vi oss av ekvation (1) (för $p = 3$),

$$c_1(\,1 \quad 1 \quad 1\,) + c_2(\,1 \quad 0 \quad 1\,) + c_2(\,0 \quad 2 \quad 1\,) = \mathbf{0}\,,$$

som vi betraktar som en homogen ekvation med obekanta c_1, c_2, c_3. Vi skall undersöka om denna ekvation har icke-triviala lösningar eller inte. Om vi skriver ut ekvationen som ett system, får vi

$$\begin{aligned} c_1 + c_2 \qquad\quad &= 0 \\ c_1 \qquad + 2c_3 &= 0\,, \\ c_1 + c_2 + \ \ c_3 &= 0 \end{aligned}$$

dvs. räkneschemat blir den matris $(\mathbf{a}_1^T \quad \mathbf{a}_2^T \quad \mathbf{a}_3^T)$, som består av de kolonner som man får då de ursprungliga radvektorerna transponeras. Allmänt gäller: **Då man undersöker det linjära beroendet och oberoendet hos vektorer i $\mathbf{R}^n$ är det naturligt och tillrådligt att alltid placera de ifrågavarande vektorerna som kolonner i en matris** oberoende av om de från början är givna som kolonnvektorer eller radvektorer. Vi överför räkneschemat på echelonform,

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

och ser att det inte finns några fria variabler. Således är den triviala lösningen $c_1 = c_2 = c_3 = 0$ den enda, vilket betyder att vektorerna $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ är linjärt oberoende.

Exempel 4.13. Vektorerna $(1 \quad 2 \quad -1)^T$, $(2 \quad -1 \quad 8)^T$ och $(-1 \quad 3 \quad -9)^T$ är linjärt beroende: Då vi bildar ett räkneschema med dessa som kolonner, får vi

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ -1 & 8 & -9 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 5 \\ 0 & 0 & 0 \end{pmatrix}.$$

Eftersom det finns en fri variabel, är vektorerna linjärt beroende. Om vi explicit vill räkna ut koefficienterna c_i $(i = 1, 2, 3)$ i det linjära beroendet, fortsätter vi till den reducerade echelonformen

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

sätter $c_3 = s$, eftersom denna variabel är fri, och får att $c_1 = -s$ och $c_2 = s$. Till slut kan vi välja $s = 1$ eller något annat värde olikt noll.

Exempel 4.14. Mängden $\{\mathbf{0}\}$ av enbart nollvektorn är linjärt beroende, ty $1 \cdot \mathbf{0} = \mathbf{0}$. Allmännare gäller att en mängd $\{\mathbf{0}, \mathbf{a}_1, \ldots, \mathbf{a}_p\}$ av vektorer, där nollvektorn ingår, är linjärt beroende, eftersom

$$1 \cdot \mathbf{0} + 0 \cdot \mathbf{a}_1 + \cdots + 0 \cdot \mathbf{a}_p = \mathbf{0}.$$

Exempel 4.15. Matriserna $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$ och $\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}$ är linjärt oberoende i vektorrummet av alla 2/2-matriser, ty ekvationen

$$c_1 \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} + c_2 \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

är ekvivalent med ekvationssystemet $c_1 + 3c_2 = 0$, $2c_2 = 0$, $2c_1 = 0$, $3c_1 + c_2 = 0$, som har den enda lösningen $c_1 = c_2 = 0$.

Exempel 4.16. Funktionerna $f_1(x) = e^x$ och $f_2(x) = x$ i vektorrummet $\mathcal{M}$ av alla funktioner definierade på $\mathbf{R}$ är linjärt oberoende: Ekvation (1), som nu har formen
$$c_1 e^x + c_2 x = 0\,,$$
skall tolkas som en likhet som gäller för varje värde på x, eftersom nollan i högerledet i detta fall representerar nollfunktionen. Likheten bör alltså gälla speciellt för t.ex. $x = 0$ och $x = 1$, vilket ger ekvationssystemet $c_1 = 0$, $ec_1 + c_2 = 0$, som har den enda lösningen $c_1 = c_2 = 0$.

Definition 4.8. Två vektorer $\mathbf{a}_1$ och $\mathbf{a}_2$ i ett vektorrum sägs vara *parallella* om de är linjärt beroende.

Om $\mathbf{a}_1$ och $\mathbf{a}_2$ är parallella, är $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$, där minst en av koefficienterna är olik noll. Om $c_1 \neq 0$, är $\mathbf{a}_1 = -\frac{c_2}{c_1}\mathbf{a}_2$ en multipel av $\mathbf{a}_2$. Om igen $c_2 \neq 0$, är $\mathbf{a}_2$ en multipel av $\mathbf{a}_1$. **Två vektorer är alltså parallella om och endast om den ena är en multipel av den andra.**

En liknande geometrisk tolkning har vi för t.ex. tre vektorer i $\mathbf{R}^3$: **Tre vektorer $\mathbf{a}_1$, $\mathbf{a}_2$ och $\mathbf{a}_3$ i $\mathbf{R}^3$ är linjärt beroende om och endast om alla befinner sig i ett och samma plan innehållande origo.** En av dem, t.ex. $\mathbf{a}_3$, är nämligen då en linjärkombination av de två andra, dvs. ligger i spannet av $\mathbf{a}_1$ och $\mathbf{a}_2$, som i sin tur är en delmängd av ett plan genom origo.

Sats 4.5. *Om $\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_k$ är vektorer i $\mathbf{R}^m$ och om $k > m$ så är dessa vektorer linjärt beroende.*

Bevis. Vi kan anta att vektorerna $\mathbf{a}_i$ är kolonnvektorer (i annat fall transponerar vi!) och sätter $A = (\mathbf{a}_1 \;\cdots\; \mathbf{a}_k)$. Då är ekvationen $x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k = \mathbf{0}$ ekvivalent med den homogena ekvationen $A\mathbf{x} = \mathbf{0}$, som enligt Sats 1.3 har icke-triviala lösningar. Satsens påstående gäller alltså. $\diamondsuit$

Om rader och kolonner i en echelonmatris

I detta avsnitt visar vi att de så kallade pivotkolonnerna och -raderna i en echelonmatris är linjärt oberoende.

Definition 4.9. *Radrummet* till en matris A är spannet av raderna i A. För radrummet använder vi beteckningen $R(A^T)$ på grund av att radrummet för A vid transponering förvandlas till kolonnrummet för A^T.

Låt först U vara en uppåt triangulär n/n-matris med n pivotelement:
$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ - & - & - & - \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}, \qquad \text{där} \quad u_{ii} \neq 0 \quad \text{för} \quad i = 1, \ldots, n\,.$$
Då är U icke-singulär, dvs. U:s kolonner $\mathbf{u}_1$, $\mathbf{u}_2$, … $\mathbf{u}_n$ är linjärt oberoende.

Vi skall visa att också U:s rader $\overline{\mathbf{u}}_1$, $\overline{\mathbf{u}}_2$, …,$\overline{\mathbf{u}}_n$ är linjärt oberoende: Den transponerade matrisen
$$U^T = (\,\overline{\mathbf{u}}_1^T \quad \overline{\mathbf{u}}_2^T \quad \ldots \quad \overline{\mathbf{u}}_n^T\,)$$

är en nedåt triangulär matris, vars diagonalelement är olika noll. Vi gör en serie framåtsubstitutioner och får att $U^T\mathbf{x} = \mathbf{0}$ bara om $\mathbf{x} = \mathbf{0}$. Följaktligen är U^T ickesingulär och kolonnerna i denna matris därför linjärt oberoende. Men då är också de transponerade kolonnerna, dvs. $\overline{\mathbf{u}}_1$, $\overline{\mathbf{u}}_2$, …,$\overline{\mathbf{u}}_n$, linjärt oberoende (en undersökning av det linjära beroendet leder ju i bägge fallen till samma ekvationssystem). Alltså gäller: **Både rader och kolonner i U är linjärt oberoende.**

Vad kan vi säga om det linjära beroendet eller oberoendet hos rader och kolonner i en echelonmatris? Låt oss ta ett numeriskt exempel:

Exempel 4.17. Betrakta echelonmatrisen

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

Vi vill utnyttja det som vi ovan kom till rörande uppåt triangulära matriser (och vill ha enklare beteckningar). Därför flyttar vi om kolonnerna så att de tre pivotkolonnerna (= de kolonner som innehåller pivotelement) kommer först:

$$U_1 = \begin{pmatrix} 1 & -1 & 3 & : & 2 & 1 \\ 0 & 1 & 2 & : & 0 & 0 \\ 0 & 0 & 4 & : & 0 & 2 \\ .. & .. & .. & : & .. & .. \\ 0 & 0 & 0 & : & 0 & 0 \end{pmatrix} \begin{array}{l} (\,\overline{\mathbf{u}}_1 \quad \overline{\mathbf{v}}_1\,) \\ (\,\overline{\mathbf{u}}_2 \quad \overline{\mathbf{v}}_2\,) \\ (\,\overline{\mathbf{u}}_3 \quad \overline{\mathbf{v}}_3\,) \\ \\ \\ \end{array}$$

Uppe i det vänstra hörnet av U_1 får vi då ett uppåt triangulärt block T. Längst till höger har vi infört beteckningar $(\,\overline{\mathbf{u}}_i \quad \overline{\mathbf{v}}_i\,)$ för de tre första raderna, där $\overline{\mathbf{u}}_i$ och $\overline{\mathbf{v}}_i$ är den del av raden som är till vänster respektive höger om den prickade linjen. På samma sätt inför vi beteckningar

$$\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{u}_3 \\ \mathbf{0} \end{pmatrix}$$

för de tre första kolonnerna i U_1. Omflyttning av kolonner påverkar inte kolonners eventuella linjära beroende eller oberoende, ty dessa egenskaper är oberoende av ordningsföljden. Inte heller den resulterande omnumreringen av radernas komponenter påverkar raders eventuella linjära beroende eller oberoende.

I det uppåt triangulära blocket T är raderna $\overline{\mathbf{u}}_i$ och kolonnerna $\mathbf{u}_i$ linjärt oberoende. Men då är också de tre första raderna och de tre första kolonnerna i U_1 linjärt oberoende, ty t.ex. ekvationen

$$c_1\,(\,\overline{\mathbf{u}}_1 \quad \overline{\mathbf{v}}_1\,) + c_2\,(\,\overline{\mathbf{u}}_2 \quad \overline{\mathbf{v}}_2\,) + c_3\,(\,\overline{\mathbf{u}}_3 \quad \overline{\mathbf{v}}_3\,) = (\,\mathbf{0} \quad \mathbf{0}\,)$$

sönderfaller i två ekvationer

$$\begin{aligned} c_1\overline{\mathbf{u}}_1 + c_2\overline{\mathbf{u}}_2 + c_3\overline{\mathbf{u}}_3 &= \mathbf{0} \\ c_1\overline{\mathbf{v}}_1 + c_2\overline{\mathbf{v}}_2 + c_3\overline{\mathbf{v}}_3 &= \mathbf{0}\,, \end{aligned}$$

av vilka redan den första ger att $c_1 = c_2 = c_3 = 0$ (en förlängning av linjärt oberoende vektorer med extra komponenter kan alltså bara göra vektorerna "mera oberoende").

Allmänt fås på detta sätt:

Sats 4.6. *Antag att U är en echelonmatris med r pivotelement. Då gäller:*

(*i*) *De r pivotkolonnerna i U är linjärt oberoende i kolonnrummet $R(U)$;*
(*ii*) *De r pivotraderna i U är linjärt oberoende i radrummet $R(U^T)$.*

Baser, koordinater och dimension

De *naturliga enhetsvektorerna*

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

i $\mathbf{R}^3$ är linjärt oberoende eftersom matrisen $(\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3) = I$ är icke-singulär. Dessutom uppspänner dessa vektorer hela $\mathbf{R}^3$, ty varje $\mathbf{x} = (x_1 \quad x_2 \quad x_3) \in \mathbf{R}^3$ är en linjärkombination

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 .$$

Dessa två egenskaper tar vi som definition på begreppet bas i ett vektorrum E:

Definition 4.10. En mängd $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ av vektorer $\mathbf{a}_i \in E$ är en *bas* i E, om

(*i*) vektorerna $\mathbf{a}_1, \mathbf{a}_2, , \ldots, \mathbf{a}_n$ är linjärt oberoende
(*ii*) och $\operatorname{spn}\{\mathbf{a}_1, \mathbf{a}_2, , \ldots, \mathbf{a}_n\} = E$.

Alternativt kan vi säga att vektorerna $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ bildar (eller utgör) en bas i E. Vektorerna $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ i basen kallas *basvektorer.*

Exempel 4.18. Vi undersöker om vektorerna

$$\mathbf{v}_1 = (1 \quad 1 \quad 2), \quad \mathbf{v}_2 = (1 \quad 2 \quad 1), \quad \mathbf{v}_3 = (3 \quad 1 \quad 1)$$

bildar en bas i $\mathbf{R}^3$. Först det linjära oberoendet: Vi sätter $V = (\mathbf{v}_1^T \quad \mathbf{v}_2^T \quad \mathbf{v}_3^T)$. Då är ekvationen $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0}$ ekvivalent med $V\mathbf{x} = \mathbf{0}$ och motsvarande kalkyler,

$$V = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & -7 \end{pmatrix},$$

visar att fria variabler saknas. Således är vektorerna linjärt oberoende.

För att påvisa att $\operatorname{spn}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathbf{R}^3$, bör vi i princip lösa $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$, dvs. $V\mathbf{x} = \mathbf{b}^T$, för varje $\mathbf{b} \in \mathbf{R}^3$. Om en lösning existerar för varje $\mathbf{b}$ så omfattar spannet av vektorerna $\mathbf{v}_i$ hela $\mathbf{R}^3$. Existensen av en lösning är emellertid garanterad, eftersom V är icke-singulär (enligt kalkylen ovan) och därmed inverterbar:

$$V\mathbf{x} = \mathbf{b}^T \iff \mathbf{x} = V^{-1}\mathbf{b}^T .$$

Den senare delen av exemplet ovan **behöver** i praktiken **aldrig utföras** på grund av följande sats, som bevisas ungefär som i exemplet:

Sats 4.7. *Om* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *är* n *stycken linjärt oberoende vektorer i* $\mathbf{R}^n$ *så bildar dessa vektorer en bas i* $\mathbf{R}^n$*.*

Bevis. Om vi sätter $A = (\,\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n\,)$, så är A en kvadratisk, icke-singulär matris. Alltså existerar A^{-1}. För varje $\mathbf{b} \in \mathbf{R}^n$ har ekvationen $A\mathbf{x} = \mathbf{b}$ då lösningen $\mathbf{x} = A^{-1}\mathbf{b}$, dvs. $\mathbf{b}$ är en linjärkombination

$$\mathbf{b} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n \in \operatorname{spn}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} .$$

Således bildar $\mathbf{a}_1, \ldots, \mathbf{a}_n$ en bas i $\mathbf{R}^n$. $\diamondsuit$

Antag att E är ett vektorrum och att $\mathbf{b}_1, \ldots, \mathbf{b}_n$ utgör en bas i E.

Definition 4.11. För ett givet $\mathbf{x} \in E$ kallas talen $x_1, \ldots, x_n$ *koordinaterna* för $\mathbf{x}$ i basen $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

Sats 4.8. *Om* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *är en bas i ett vektorrum* E*, så är koordinaterna för ett givet* $\mathbf{x} \in E$ *entydiga.*

Bevis. Antag att både $x_1, \ldots, x_n$ och $y_1, \ldots, y_n$ är koordinater för ett givet $\mathbf{x}$ i den givna basen. Då är

$$\begin{aligned} \mathbf{x} &= x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \\ \mathbf{x} &= y_1\mathbf{b}_1 + \cdots + y_n\mathbf{b}_n , \end{aligned}$$

vilket leder till att

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (x_1 - y_1)\mathbf{b}_1 + \cdot + (x_n - y_n)\mathbf{b}_n .$$

Eftersom basvektorerna är linjärt oberoende, är alla koefficienterna i linjärkombinationen i högerledet noll, dvs. $x_1 = y_1, \ldots, x_n = y_n$. $\diamondsuit$

Exempel 4.19. Antag att vi fått som uppgift att både visa att vektorerna

$$\mathbf{b}_1 = (\,1 \quad 1 \quad 1\,)^T , \quad \mathbf{b}_2 = (\,0 \quad 2 \quad 1\,)^T , \quad \mathbf{b}_3 = (\,1 \quad 0 \quad 2\,)^T$$

bildar en bas i $\mathbf{R}^3$ och att bestämma koordinaterna för $\mathbf{a} = (\,5 \quad 5 \quad 8\,)^T$ i denna bas. Då kan dessa två uppgifter lösas **samtidigt** på följande sätt: Vi löser det ekvationssystem som svarar mot räkneschemat $(\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad | \quad \mathbf{a}\,)$ och ser redan då vi har nått någon echelonform,

$$(\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad | \quad \mathbf{a}\,) = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 1 & 2 & 0 & 5 \\ 1 & 1 & 2 & 8 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & -3 & -6 \end{array}\right) ,$$

att inga variabler är fria, varför vektorerna $\mathbf{b}_i$ är linjärt oberoende. Eftersom vi nu har tre linjärt oberoende vektorer i $\mathbf{R}^3$ så bilder dessa en bas i $\mathbf{R}^3$ enligt Sats 4.7. För att få fram koordinaterna för $\mathbf{a}$, fortsätter vi tills vi har en reducerad echelonmatris och kan då avläsa dessa: $x_1 = 3$, $x_2 = 1$, $x_3 = 2$.

Exempel 4.20. Vi skall visa att polynomen

$$1, \quad x, \quad x^2, \dots, \quad x^n$$

bildar en bas i vektorrummet P_n av alla polynom av högst graden n. Eftersom varje polynom p i P_n har formen

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n ,$$

är det klart att de givna polynomen uppspänner hela P_n. Vi behöver alltså bara visa att de också är linjärt oberoende. För detta ändamål sätter vi en linjärkombination av dem lika med nollfunktionen:

$$c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n = 0 .$$

Denna likhet gäller för varje x, alltså speciellt också för $x = 0$, vilket ger att $c_0 = 0$. Således är

$$x(c_1 + c_2 x + \cdots + c_n x^{n-1}) = 0$$

för varje x, dvs.

$$c_1 + c_2 x + \cdots + c_n x^{n-1} = 0$$

för varje $x \neq 0$. Men polynom är kontinuerliga, så polynomet $c_1 + c_2 x + \cdots + c_n x^{n-1}$ i vänsterledet måste vara noll också för $x = 0$, vilket ger att $c_1 = 0$ osv. Ett antal upprepningar av detta resonemang ger att $c_0 = c_1 = \cdots = c_n = 0$ (strängt taget borde man använda induktion!). De givna polynomen bildar alltså en bas i P_n.

Nästa sats visar att antalet vektorer i en bas är entydigt:

Sats 4.9. *Om både vektorerna $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_m$ och vektorerna $\mathbf{b}_1$, $\mathbf{b}_2$, ..., $\mathbf{b}_n$ bildar baser i ett vektorrum E, så är $m = n$.*

Bevis. Antag som en **antites** att det motsatta gäller, t.ex. att $m < n$ (om $m > n$ kan man låta de två baserna byta plats i beviset). Vi skall visa att detta antagande leder till en motsägelse. Varje vektor $\mathbf{b}_i$ kan skrivas som en linjärkombination av vektorerna $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_m$:

$$\mathbf{b}_i = \sum_{k=1}^{m} \beta_{ki} \mathbf{a}_k \quad (i = 1, \dots, n) .$$

Sätt $B = (\beta_{ki})$. Då har det homogena systemet $B\mathbf{x} = \mathbf{0}$ icke-triviala lösningar enligt Sats 1.3, ty antalet rader m är mindre än antalet obekanta n. Men å andra sidan kan vi bevisa det motsatta genom att sluta oss till att systemet bara kan ha den triviala lösningen:

$$\begin{aligned}
B\mathbf{x} = \mathbf{0} \quad &\Longrightarrow \quad \sum_{i=1}^{n} \beta_{ki} x_i = 0 \quad (k = 1, \ldots, m) \\
&\Longrightarrow \quad \mathbf{0} = \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \beta_{ki} x_i \right) \mathbf{a}_k = \sum_{i=1}^{n} x_i \left(\sum_{k=1}^{m} \beta_{ki} \mathbf{a}_k \right) = \sum_{i=1}^{n} x_i \mathbf{b}_i \\
&\Longrightarrow \quad x_1 = \ldots = x_n = 0 \,,
\end{aligned}$$

där vi för den sista implikationen har utnyttjat att $\mathbf{b}_1, \ldots, \mathbf{b}_n$ är linjärt oberoende. Men detta är en motsägelse. På grund av detta måste vi dra slutsatsen att $m = n$. $\diamondsuit$

Antalet vektorer i en bas är enligt Sats 4.9 ett tal som karakteriserar ett visst vektorrum. Vi inför därför följande definition:

Definition 4.12. Med *dimensionen*, $\dim E$, hos ett vektorrum E menas antalet vektorer i en bas i E.

Anmärkning. Man kan visa att varje vektorrum $E \neq \{\mathbf{0}\}$ har en bas, ändlig eller oändlig. Oändliga baser definieras på samma sätt som ändliga. Skillnaden är bara att med en linjärkombination av en oändlig mängd av vektorer avses en linjärkombination av någon ändlig delmängd av denna. Om ett vektorrum E har en oändlig bas, säger man att E är oändligtdimensionellt och skriver att $\dim E = \infty$.

Exempel 4.22. Tydligen är

$$\dim \mathbf{R}^n = n \,,$$

eftersom de naturliga enhetsvektorerna $\mathbf{e}_1, \ldots, \mathbf{e}_n$ bildar en bas i $\mathbf{R}^n$, och

$$\dim P_n = n + 1 \,,$$

eftersom polynomen $1, x, x^2, \ldots, x^n$ bildar en bas i P_n. Vektorrummet P av alla polynom, liksom vektorrummet av alla funktioner, måste vara oändligtdimensionellt, eftersom den oändliga mängden av alla potenser x^n av x är linjärt oberoende.

I beviset av Sats 4.9 utnyttjade vi inte alla antaganden i satsen efter att antitesen var gjord: Vi använde oss bara av att $\mathbf{a}_1, \ldots, \mathbf{a}_m$ uppspänner E och att $\mathbf{b}_1, \ldots, \mathbf{b}_n$ är linjärt oberoende. Precis samma slag av resonemang ger oss därför följande generalisering av Sats 4.7:

Sats 4.10. *Om $\mathbf{b}_1, \ldots, \mathbf{b}_m$ är linjärt oberoende vektorer i ett vektorrum E och $\dim E = m$, så bildar $\mathbf{b}_1, \ldots, \mathbf{b}_m$ en bas i E.*

Bevis. Antag som en antites att $\operatorname{spn}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\} \neq E$ så att vi kan välja en vektor $\mathbf{b}_{m+1} \in E$ utanför det nämnda spannet. Då är $\mathbf{b}_1, \ldots, \mathbf{b}_{m+1}$ linjärt oberoende (enligt uppgift 14 (a)). Beviset för föregående sats (med $n = m + 1$) ger nu en motsägelse. Alltså är $\operatorname{spn}\{\mathbf{b}_1, \ldots, \mathbf{b}_m\} = E$, dvs. $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ är en bas. $\diamondsuit$

Exempel 4.21. Vi skall visa att polynomen

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

bildar en bas i vektorrummet P_2 av polynom av högst graden 2. Eftersom vi redan vet att de tre polynomen 1, x och x^2 bildar en bas, räcker det enligt Sats 4.10 att visa att p_1, p_2 och p_3 är linjärt oberoende. Vi undersöker därför ekvationen

$$\begin{aligned} 0 &= c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x) \\ &= (c_1 + c_2 + 2c_3) + (c_1 + 2c_2 + c_3)x + (c_1 + c_2)x^2 , \end{aligned}$$

som gäller för varje x om och endast om ekvationssystemet

$$\begin{aligned} c_1 + \ \ c_2 + 2c_3 &= 0 \\ c_1 + 2c_2 + \ \ c_3 &= 0 \\ c_1 + \ \ c_2 \qquad &= 0 \end{aligned}$$

är satisfierat. En kalkyl med motsvarande räkneschema visar att den triviala lösningen är den enda. Således bildar p_1, p_2 och p_3 en bas i P_2.

Bas och dimension för kolonn- och radrum

Låt oss återkalla i minnet att om A är en matris så är kolonnrummet $R(A)$ spannet av A:s kolonner, radrummet $R(A^T)$ spannet av A:s rader samt nollrummet $N(A)$ vektorrummet av alla lösningar till $A\mathbf{x} = \mathbf{0}$.

Vi kommer nu att härleda algoritmer som ger baser i $R(A)$, $R(A^T)$ och $N(A)$. För kolonn- och radrummens del kommer vi att stöda oss på Sats 4.6. Vi börjar med den algoritm, som ger en bas i $R(A^T)$:

Sats 4.11. *Antag att A är en m/n-matris och antag att vi med hjälp av basoperationer har transformerat A till någon echelonform U. Då gäller:*

(*i*) *Matriserna A och U har samma radrum;*

(*ii*) *Pivotraderna i U bildar en bas i radrummet $R(A^T)$.*

Bevis. (*i*) Låt

$$X = \begin{pmatrix} \overline{\mathbf{x}}_1 \\ \vdots \\ \overline{\mathbf{x}}_m \end{pmatrix} \quad \overset{BO}{\longrightarrow} \quad Y = \begin{pmatrix} \overline{\mathbf{y}}_1 \\ \vdots \\ \overline{\mathbf{y}}_m \end{pmatrix}$$

vara en godtycklig användning av en basoperation i kedjan $A \to \cdots \to U$. Då är varje $\overline{\mathbf{y}}_j$ en linjärkombination av vektorerna $\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_m$, ty varje användning av en basoperation innebär att man bildar en (enkel) linjärkombination av rader. Eftersom alla basoperationer är omvändbara, gäller också det omvända: Varje vektor $\overline{\mathbf{x}}_j$ är en linjärkombination av $\overline{\mathbf{y}}_1, \ldots, \overline{\mathbf{y}}_m$. Ur detta följer att

$$\operatorname{spn}\{\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_m\} = \operatorname{spn}\{\overline{\mathbf{y}}_1, \ldots, \overline{\mathbf{y}}_m\} ,$$

dvs. att X och Y har samma radrum. Då detta gäller för varje länk i kedjan från A till U, är $R(A^T) = R(U^T)$.

(ii) Det är klart att pivotraderna i U bildar en bas i $R(U^T) = R(A^T)$, ty de är linjärt oberoende enligt Sats 4.6 och uppspänner hela radrummet, eftersom de övriga raderna i U är $\mathbf{0}$. $\diamondsuit$

Exempel 4.23. Om vi överför matrisen A nedan till en echelonmatris,

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 1 & -1 \\ 0 & 2 & -8 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{pmatrix} = U\,,$$

så ser vi att U har pivotelement i de två första raderna, som alltså bildar en bas

$$\{(1 \quad 1 \quad -1)\,,\,(0 \quad -1 \quad 4)\}$$

i $R(A^T)$. Observera att vi kan använda vilken echelonform som helst. Om vi använder t.ex. den reducerade echelonformen, hittar vi en annan bas i $R(A^T)$.

Lägg märke till att vektorerna skall tas ur U, inte ur A. Vi ger ett motexempel:

Exempel 4.24. Sats 4.11 ger på basen av kalkylen

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = U\,,$$

att den första raden i U bildar en bas $\{(1 \quad 0)\}$ i $R(A^T)$. Däremot bildar inte den första raden i A någon bas i $R(A^T)$, ty den är ju linjärt beroende.

För kolonnrummet $R(A)$ har vi en något annorlunda algoritm för konstruktion av en bas. Orsaken är att matrisen A och en motsvarande echelonmatris U i regel har olika kolonnrum. Detta i sin tur sammanhänger med att basoperationerna är radoperationer, inte kolonnoperationer.

Sats 4.12. *Låt A vara en m/n-matris och antag att vi med hjälp av basoperationerna transformerat A till någon echelonmatris U. Då bilder de kolonner i A, som har* samma nummer *som pivotkolonnerna i U, en bas i kolonnrummet $R(A)$.*

Bevis. Vi inför beteckningar för kolonnerna i A och U:

$$\begin{aligned} A &= (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n)\,, \\ U &= (\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_n)\,. \end{aligned}$$

Echelonmatrisen U fås genom multiplikation från vänster med (i tur och ordning) matriser E_1, E_2, …, E_p, som svarar mot basoperationer. Om vi sätter $E = E_p \cdots E_1$ så är E inverterbar och

$$\begin{aligned} U &= EA\,, \\ \mathbf{u}_i &= E\mathbf{a}_i \quad (i = 1, \ldots, n)\,. \end{aligned} \tag{3}$$

Antag att U innehåller r pivotkolonner. Vi kan anta att de r första, $\mathbf{u}_1, \ldots, \mathbf{u}_r$, råkar vara pivotkolonner. Om inte så kan vi ju temporärt numrera om kolonnerna!

$$
\begin{array}{cc}
 & \begin{array}{ccccccc} \mathbf{u}_1 & \mathbf{u}_2 & & \mathbf{u}_r & & & \end{array} \\
U = & \begin{pmatrix}
P & * & .. & * & * & .. & * \\
 & P & .. & * & * & .. & * \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
 & & & P & * & .. & * \\
0 & 0 & .. & 0 & 0 & .. & 0 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots \\
0 & 0 & .. & 0 & 0 & .. & 0
\end{pmatrix} .
\end{array}
$$

(Här betecknar P ett pivotelement). Då är dessa r pivotkolonner linjärt oberoende enligt Sats 4.6. Vilken som helst av de övriga kolonnerna, säg $\mathbf{u}_k$ $(k > r)$, är en linjärkombination av $\mathbf{u}_1, \ldots, \mathbf{u}_r$, ty $\mathbf{u}_k$ kan betraktas som ett högerled i ett räkneschema $(\mathbf{u}_1 \ \cdots \ \mathbf{u}_r \mid \mathbf{u}_k)$ och bakåtsubstitution ger då koefficienterna i linjärkombinationen

$$x_1\mathbf{u}_1 + \cdots + x_r\mathbf{u}_r = \mathbf{u}_k \,.$$

Alltså är $R(U) = \operatorname{spn}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\} = \operatorname{spn}\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$, dvs. $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ är en bas i $R(U)$.

Då vi nu vet att pivotkolonnerna i U bildar en bas i $R(U)$, kan vi använda (3) till att visa att motsvarande kolonner i A, dvs. $\mathbf{a}_1, \ldots, \mathbf{a}_r$, bildar en bas i $R(A)$. Först det linjära oberoendet: Om vi multiplicerar ekvationen

$$c_1\mathbf{a}_1 + \cdots + c_r\mathbf{a}_r = \mathbf{0}$$

från vänster med matrisen E och beaktar (3), får vi

$$\mathbf{0} = c_1E\mathbf{a}_1 + \cdots + c_rE\mathbf{a}_r = c_1\mathbf{u}_1 + \cdots + c_r\mathbf{u}_r \,,$$

vilket ger att $c_1 = \cdots = c_r = 0$ då ju $\mathbf{u}_1, \ldots, \mathbf{u}_r$ är linjärt oberoende. För att visa att spannet av $\mathbf{a}_1, \ldots, \mathbf{a}_r$ är hela $R(A)$, tar vi en godtycklig vektor $\mathbf{x} \in R(A)$. Då är $E\mathbf{x} \in R(U)$, enligt (3), och kan därför skrivas som en linjärkombination

$$E\mathbf{x} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r \,.$$

Således är $\mathbf{x} = \alpha_1E^{-1}\mathbf{u}_1 + \cdots + \alpha_rE^{-1}\mathbf{u}_r = \alpha_1\mathbf{a}_1 + \cdots + \alpha_r\mathbf{a}_r$. Varje vektor $\mathbf{x} \in R(A)$ finns alltså i $\operatorname{spn}\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ och satsen är därmed bevisad. $\diamondsuit$

Definition 4.13. *Rangen* hos en matris A är dimensionen $\dim R(A)$ hos kolonnrummet. Dimensionen hos nollrummet, $\dim N(A)$, sägs vara A:s *defekt* (eller nolldefekt).

Enligt satserna 4.11 och 4.12 är

$$\dim R(A) = \dim R(U) = \#(\text{pivotelement}) = \dim R(A^T) \,.$$

Följaktligen gäller:

Sats 4.13. *A och A^T har samma rang för varje matris A.*

Med hjälp av detta kan vi skriva ut en stor mängd påståenden, som alla är ekvivalenta med att en n/n-matris A är inverterbar:

$$\begin{array}{ccccc}
A \text{ inverterbar} & \Leftrightarrow & A \text{ icke-singulär} & \Leftrightarrow & \dim N(A) = 0 \\
 & & \Updownarrow & & \\
 & & \dim R(A) = n & & \\
 & & \Updownarrow & & \\
 & & \dim R(A^T) = n & & \\
 & & \Updownarrow & & \\
A^T \text{ inverterbar} & \Leftrightarrow & A^T \text{ icke-singulär} & \Leftrightarrow & \dim N(A^T) = 0 \,.
\end{array}$$

Exempel 4.25. Som ett exempel på hur man använder Sats 4.12 betraktar vi en matris A och en echelonform U av denna:

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = U\,.$$

Eftersom den första och den tredje kolonnen i U är pivotkolonner, bildar den första och den tredje kolonnen i A en bas i $R(A)$:

$$\{(1 \quad 2 \quad -1)^T, (3 \quad 9 \quad 3)^T\}\,.$$

Anmärkning. (i) Man kan naturligtvis använda (t.ex.) proceduren för att bestämma en bas i ett radrum till att bestämma en bas i $R(A)$ genom att helt enkelt tillämpa den på matrisen A^T i stället för A.

(ii) En bas i ett vektorrum är ju alltid precis detsamma som en **maximal mängd av linjärt oberoende vektorer** i ett vektorrum. Om man bland ett antal givna vektorer i $\mathbf{R}^n$ vill välja en bas i spannet av dem, skriver man därför vektorerna som **kolonner** i en matris A och använder Sats 4.12 till att bestämma en bas i $R(A)$, som ju är spannet av de givna vektorerna.

Sambandet mellan rang och defekt

Först kan vi konstatera att den procedur som vi använde i exempel 4.9 för att skriva ett nollrum som ett spann, i själva verket alltid ger en bas i nollrummet:

Exempel 4.26. Matrisen A överförs i reducerad echelonform

$$\begin{pmatrix} 1 & 3 & 3 & -1 \\ 2 & 6 & 7 & 2 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 3 & 0 & -13 \\ 0 & 0 & 1 & 4 \end{pmatrix},$$

de fria variablerna ges godtyckliga värden, $x_2 = s$ och $x_4 = t$, varefter lösningen först skrivs ut komponentvis

$$\begin{aligned} x_1 &= -3s + 13t \\ x_2 &= s \\ x_3 &= - 4t \\ x_4 &= t \end{aligned}$$

och sedan i vektorform

$$\mathbf{x} = s\begin{pmatrix} -3 \\ 1^* \\ 0 \\ 0^* \end{pmatrix} + t\begin{pmatrix} 13 \\ 0^* \\ -4 \\ 1^* \end{pmatrix} = s\mathbf{v}_1 + t\mathbf{v}_2 .$$

De komponenter som svarar mot de fria variablerna har här försetts med en stjärna. Nollrummet är nu spannet av $\mathbf{v}_1$ och $\mathbf{v}_2$ men dessa vektorer är också linjärt oberoende, ty ekvationen $s\mathbf{v}_1 + t\mathbf{v}_2 = \mathbf{0}$ ger direkt att $s = t = 0$ då vi utnyttjar de stjärnförsedda komponenterna som svarar mot fria variabler. Vektorerna $\mathbf{v}_1$ och $\mathbf{v}_2$ bildar alltså en **bas** i $N(A)$.

Generellt gäller precis som i vårt exempel, att $\dim N(A) = \#(\text{fria variabler})$ så att

$$\begin{array}{r} \dim R(A) = \#(\text{pivotkolonner}) \\ \dim N(A) = \#(\text{icke-pivotkolonner}) \\ \hline \dim R(A) + \dim N(A) = \#(\text{kolonner})\,. \end{array}$$

Sats 4.14. *Antag att A är en m/n-matris. Då gäller:*

(*i*) $\dim R(A) + \dim N(A) = n\,;$

(*ii*) *En bas $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ i nollrummet $N(A)$ fås genom att man i lösningen till $A\mathbf{x} = \mathbf{0}$,*

$$\mathbf{x} = s_1\mathbf{v}_1 + \cdots + s_p\mathbf{v}_p\,,$$

i tur och ordning sätter en *fri variabel s_i lika med* 1 *medan de övriga får värdet* 0.

Om man enbart skall bestämma defekten hos en matris A, räcker det att överföra A i echelonform. Antalet fria variabler anger då antalet vektorer i en bas i $N(A)$, dvs. defekten. Antalet basvariabler anger på samma sätt matrisens rang.

Exempel 4.27. För matrisen A nedan, fås en echelonform med två fria variabler och två basvariabler:

$$A = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 2 & 3 & 1 \\ 1 & 5 & 10 & 3 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 3 & 7 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

Således är både defekten och rangen likamed 2.

Vänsternollrum. Matriser med rangen ett

Vi har redan sett att kolonnrummet $R(A^T)$ till den transponerade matrisen A^T på ett naturligt sätt (via transponering) kan identifieras med radrummet till A. Då det gäller nollrummet till A^T, får vi genom transponering:

$$A^T\mathbf{x} = \mathbf{0} \quad \Leftrightarrow \quad (A^T\mathbf{x})^T = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{x}^T A = \mathbf{0}\,.$$

Detta ger anledning till följande benämning:

Definition 4.14. *Vänsternollrummet* till en m/n-matris A är mängden

$$N(A^T) = \{\mathbf{x} \mid \mathbf{x}^T A = \mathbf{0}\}\,.$$

Denna är identisk med nollrummet till A^T och är alltså ett underrum av $\mathbf{R}^m$.

Med varje m/n-matris har vi nu associerat fyra underrum, kolonnrummet, radrummet, nollrummet och vänsternollrummet, av vilka det sistnämnda är det minst viktiga. Om A har rangen r så gäller sammanfattningsvis:

$R(A)$	A:s kolonnrum;	dimensionen $= r$
$R(A^T)$	A:s radrum;	dimensionen $= r$
$N(A)$	A:s nollrum;	dimensionen $= n - r$
$N(A^T)$	A:s vänsternollrum;	dimensionen $= m - r$

Vi skall nu undersöka "strukturen" hos en m/n-matris

$$A = (\,\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n\,)$$

med rangen 1. Någon av kolonnerna, t.ex. $\mathbf{b} = \mathbf{a}_k \neq \mathbf{0}$, bildar en bas i $R(A)$. Varje $\mathbf{a}_i$ är alltså en linjärkombination $\mathbf{a}_i = c_i\mathbf{b}$ av $\mathbf{b}$, dvs.

$$A = (\,c_1\mathbf{b} \quad \cdots \quad c_n\mathbf{b}\,) = \mathbf{b}\,(\,c_1 \quad \cdots \quad c_n\,) = \mathbf{b}\mathbf{c}\,,$$

där $c_i \neq 0$ för minst ett i (annars är ju A:s rang noll). Omvänt: Om $A = \mathbf{b}\mathbf{c}$, där $\mathbf{b} \neq \mathbf{0}$, är en kolonnvektor och $\mathbf{c} \neq \mathbf{0}$ är en radvektor är $A = (\,c_1\mathbf{b} \quad \cdots \quad c_n\mathbf{b}\,)$. Men då uppspänns ju $R(A)$ av $\mathbf{b}$ så att $\dim R(A) = 1$.

En matris A har alltså rangen 1 om och endast om $A = \mathbf{b}\mathbf{c}$, där $\mathbf{b} \neq \mathbf{0}$ är en kolonnvektor och $\mathbf{c} \neq \mathbf{0}$ är en radvektor.

Ett liknande resonemang ger att en matris A har rangen 2 om och endast om den har formen $A = \mathbf{b}_1\mathbf{c}_1 + \mathbf{b}_2\mathbf{c}_2$, där $\mathbf{b}_1$, $\mathbf{b}_2$ är linjärt oberoende kolonnvektorer och $\mathbf{c}_1$, $\mathbf{c}_2$ är linjärt oberoende radvektorer (och A har rangen 3 om och endast ...)

Exempel 4.28. I matrisen

$$A = \begin{pmatrix} 6 & 3 & 3 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{pmatrix},$$

som har rangen 1, bildar (t.ex.) första kolonnen $\mathbf{b}$ en bas i $R(A)$ så att

$$A = (\,1\mathbf{b} \quad (1/2)\mathbf{b} \quad (1/2)\mathbf{b}\,) = \mathbf{b}\mathbf{c}\,,$$

där $\mathbf{c} = (\,1 \quad 1/2 \quad 1/2\,)$.

Höger- och vänsterinverser

Antag att A är en m/n-matris. För A:s rang $r(A)$ gäller både $r(A) = \dim R(A) \le n$ och $r(A) = \dim R(A^T) \le m$, dvs. om $\min(m,n)$ är det mindre av talen m och n, är

$$r(A) \le \min(m,n)\,. \tag{4}$$

Vi kommer att visa att likhet gäller i denna formel om och endast om matrisen A har en så kallad höger- eller en vänsterinvers. Sådana **ensidiga** inverser definieras på följande sätt:

Definition 4.15. Om $AB = I$ så är B en *högerinvers* till A och A är en *vänsterinvers* till B.

Vi ser av denna definition, att inversen till en kvadratisk matris A är en matris som är både höger- och vänsterinvers till A.

Villkoret för existensen av en högerinvers följer ur nedanstående resonemang, där vi har använt oss av beteckningar för kolonnerna i en (obekant) högerinvers $X = (\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_m\,)$ och enhetsmatrisen $I = (\,\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_m\,)$:

$$\begin{aligned}
AX = I_m \text{ är lösbar} \quad &\Leftrightarrow \quad A\,(\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_m\,) = (\,\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_m\,) \text{ är lösbar}\\
&\Leftrightarrow \quad A\mathbf{x}_i = \mathbf{e}_i \text{ är lösbar } (i = 1, \ldots, m)\\
&\Leftrightarrow \quad \mathbf{e}_i \in R(A) \quad (i = 1, \ldots, m)\\
&\Leftrightarrow \quad R(A) = \mathbf{R}^m\\
&\Leftrightarrow \quad r(A) = \dim R(A) = m\,.
\end{aligned}$$

Ur detta kan vi utläsa nedanstående sats:

Sats 4.15. *För en m/n-matris A är följande påståenden ekvivalenta:*

(*i*) *Ekvationen $A\mathbf{x} = \mathbf{b}$ har minst* en *lösning för varje $\mathbf{b} \in \mathbf{R}^m$;*
(*ii*) $R(A) = \mathbf{R}^m$;
(*iii*) $r(A) = \dim R(A) = m$;
(*iv*) *A har en högerinvers.*

Observera att enligt (4) är det möjligt att uppfylla (*iii*) bara om $m \le n$.

För att räkna ut en högerivers till A löser man matrisekvationen $AX = I$ med hjälp av räkneschemat $(A \mid I)$. Detta svarar, som vi har sett i kapitel 3, mot samtidig lösning av m ekvationssystem med samma koefficientmatris A:

Exempel 4.29. Om A är av typen $2/3$ så har en eventuell högerinvers X typen $3/2$, så att om

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -3 & 3 \end{pmatrix} \quad \text{så är} \quad X = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix}$$

och de obekanta elementen i X fås ur en reducerad echelonmatris:

$$(A \mid I) = \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ -1 & -3 & 3 & 0 & 1 \end{array}\right) \to \cdots \to \left(\begin{array}{ccc|cc} 1 & 0 & 15 & 3 & 2 \\ 0 & 1 & -6 & -1 & -1 \end{array}\right).$$

Eftersom x_3 och y_3 är fria, sätter vi $x_3 = s$ och $y_3 = t$ och får variablerna x_i genom att beakta den första kolonnen i högerledet och variablerna y_i genom att beakta den andra kolonnen i högerledet:

$$X = \begin{pmatrix} 3-15s & 2-15t \\ -1+6s & -1+6t \\ s & t \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & -1 \\ 0 & 0 \end{pmatrix} + s\begin{pmatrix} -15 & 0 \\ 6 & 0 \\ 1 & 0 \end{pmatrix} + t\begin{pmatrix} 0 & -15 \\ 0 & 6 \\ 0 & 1 \end{pmatrix}.$$

För varje insatt värde på s och t är X en högerinvers och varje högerinvers har denna form.

Vänsterinverser Y till A fås genom att man löser matrisekvationen $YA = I$, som är ekvivalent med $A^TY^T = I$. Vänsterinverser till A svarar alltså (via transponering) mot högerinverser till A^T och tvärtom. Enligt Sats 4.15 och Sats 4.13 har $A^TY^T = I$ en lösning Y^T om och endast om $\dim R(A) = \dim R(A^T) = n$, ty A^T är en n/m-matris. Detta ingår i beviset av nästa sats:

Sats 4.16. *För en m/n-matris A är följande påståenden ekvivalenta:*

- (*i*) *Ekvationen $A\mathbf{x} = \mathbf{b}$ har högst* en *lösning för varje $\mathbf{b} \in \mathbf{R}^m$;*
- (*ii*) $r(A) = \dim R(A) = n$;
- (*iii*) *Kolonnerna i A är alla linjärt oberoende;*
- (*iv*) *A har en vänsterinvers.*

Observera att enligt (4) är det möjligt att uppfylla (*ii*) bara om $m \geq n$.

Bevis. Vi har just visat att (*ii*) är ekvivalent med (*iv*) och dessutom är det uppenbart att (*ii*) är ekvivalent med (*iii*). Det räcker alltså att visa att (*i*) är ekvivalent med (*ii*):

Enligt Sats 4.4 har alla lösningar till en ekvation $A\mathbf{x} = \mathbf{b}$ formen $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$, där $\mathbf{x}_0$ är en partikulärlösning till denna ekvation och $\mathbf{y}$ är någon lösning till motsvarande homogena ekvation $A\mathbf{y} = \mathbf{0}$. Ur detta följer:

$$\begin{aligned} A\mathbf{x} = \mathbf{b} \text{ har högst en lösning} \quad &\Leftrightarrow \quad A\mathbf{y} = \mathbf{0} \text{ bara om } \mathbf{y} = \mathbf{0} \\ &\Leftrightarrow \quad \#(\text{fria variabler}) = 0 \\ &\Leftrightarrow \quad \dim R(A) = n\,. \end{aligned}$$

Således är (*i*) ekvivalent med (*ii*). $\diamondsuit$

Som vi redan nämnde, så räknar man ut vänsterinverser till en matris A genom att man löser ekvationen $YA = I$ eller rättare sagt det transponerade systemet $A^TY^T = I$ med hjälp av räkneschemat $(A^T \mid I)$, vilket ger Y^T och slutligen Y:

Exempel 4.30. Om A är av typen $3/2$ så bör en eventuell vänsterinvers Y ha typen $2/3$. Detta betyder att om

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{och om vi sätter} \quad Y = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

så ger kalkylerna

$$(A^T \mid I) = \begin{pmatrix} 1 & 2 & 3 & | & 1 & 0 \\ 1 & 2 & 4 & | & 0 & 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 2 & 0 & | & 4 & -3 \\ 0 & 0 & 1 & | & -1 & 1 \end{pmatrix},$$

där vi ser att x_2 och y_2 är fria, vänsterinverserna

$$Y = \begin{pmatrix} 4-2s & s & -1 \\ -3-2t & t & 1 \end{pmatrix} \quad (s,\, t \in \mathbf{R}).$$

Med hjälp av den teori som vi nu har utvecklat, kan vi bevisa att om en kvadratisk matris har en högerinvers så har den också en vänsterinvers och att dessa inverser då är lika:

Sats 4.17. *Antag att A är en m/n-matris och låt A_H och A_V beteckna någon höger- respektive vänsterinvers till A. Då gäller:*

(*i*) *Om både ett A_H och ett A_V existerar, så är $m = n$.*

(*ii*) *Om $m = n$, så existerar ett A_H om och endast om ett A_V existerar och då är*

$$A_H = A_V\,,$$

dvs. A är inverterbar och $A^{-1} = A_H = A_V$.

Bevis. Med hjälp av satserna 4.15 och 4.16 fås

$$\left.\begin{array}{lcl} A_H \text{ existerar} & \Leftrightarrow & \dim R(A) = m \\ A_V \text{ existerar} & \Leftrightarrow & \dim R(A) = n \end{array}\right\} \quad \Rightarrow \quad m = n\,.$$

Om $m = n$ och och A_H och A_V existerar, så är $A_H = (A_V A)A_H = A_V(AA_H) = A_V$, eftersom både $A_V A$ och AA_H är enhetsmatriser. $\diamondsuit$

Övningsuppgifter

1. Visa, att $E = \{(\,x_1 \quad x_2\,) \mid x_1, x_2 \in \mathbf{R}, x_1 + x_2 = 2\}$ med räkneoperationerna definierade genom

$$\begin{aligned} (\,x_1 \quad x_2\,) + (\,y_1 \quad y_2\,) &= (\,x_1 + y_1 - 1 \quad x_2 + y_2 - 1\,)\,, \\ \lambda\,(\,x_1 \quad x_2\,) &= (\,\lambda x_1 + 1 - \lambda \quad \lambda x_2 + 1 - \lambda\,) \end{aligned}$$

är ett vektorrum.

2. Vilka av följande delmängder av $\mathbf{R}^3$ är underrum då räkneoperationerna är de vanliga:

(*a*) Mängden av vektorer med första komponenten 0?

(*b*) Mängden av vektorer med första komponenten 1?

(*c*) $\{(\,0 \quad 0 \quad 0\,)\}$?

(*d*) $\{(\,b_1 \quad b_2 \quad b_3\,) \mid 3b_1 - b_2 + b_3 = 0\}$?

3. Mängden $\mathbf{R}^\infty = \{(\,x_1 \quad x_2 \quad \cdots\,) \mid x_j \in \mathbf{R},\ j = 1, 2, \ldots\}$ är ett vektorrum om

$$\begin{aligned}(\,x_1 \quad x_2 \quad \cdots\,) + (\,y_1 \quad y_2 \quad \cdots\,) &= (\,x_1 + y_1 \quad x_2 + y_2 \quad \cdots\,)\,,\\ \lambda\,(\,x_1 \quad x_2 \quad \cdots\,) &= (\,\lambda x_1 \quad \lambda x_2 \quad \cdots\,)\,.\end{aligned}$$

Vilka av följande delmängder av $\mathbf{R}^\infty$ är underrum:

(a) Mängden av alla följder som innehåller oändligt många nollor (t.ex. $(\,1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad \cdots\,)$)?

(b) Alla följder $(\,x_1 \quad x_2 \quad \cdots\,)$ med $x_j = 0$ från och med något index $j = j_0$ (t.ex. $(\,1 \quad 2 \quad 3 \quad 1 \quad 0 \quad 0 \quad 0 \quad \cdots\,)$)?

(c) Alla aritmetiska följder, dvs. sådana för vilka $x_{j+1} - x_j$ har samma värde för varje j?

4. (a) Verifiera att mängden $\mathcal{M}$ av alla funktioner $f : \mathbf{R} \to \mathbf{R}$ med definitionsmängden $[a,\, b]$ är ett vektorrum, då räkneoperationerna definieras som i exempel 4.2.

   (b) Verifiera att mängden P_n av alla polynom $p(x)$ av högst graden n är ett vektorrum genom att visa att det är ett underrum av $\mathcal{M}$.

5. Lös ekvationen

$$(\,0 \quad 2\,) + 3\,(\,t_1 \quad t_2\,) = (\,3 \quad -1\,)$$

   i vektorrummet i uppgift 1.

6. Visa att mängden av de positiva reella talen blir ett vektorrum då räkneoperationerna definieras enligt: Vektoradditionen $x \oplus y$ är vanlig multiplikation xy; multiplikationen $\lambda \top x$ av en skalär λ med en vektor x är potensen x^λ.

7. Undersök om vektorerna

$$\begin{aligned}&(a)\quad (\,3 \quad 2 \quad 5 \quad 1\,),\, (\,2 \quad 3 \quad 3 \quad 2\,),\, (\,1 \quad -1 \quad 2 \quad -1\,),\, (\,1 \quad 1 \quad 1 \quad 1\,)\,,\\ &(b)\quad (\,2 \quad 2 \quad 3 \quad 3\,),\, (\,1 \quad -1 \quad -1 \quad 2\,),\, (\,3 \quad 1 \quad 2 \quad 5\,),\, (\,3 \quad 3 \quad 3 \quad 3\,)\,,\end{aligned}$$

   i $\mathbf{R}^4$ är linjärt beroende eller oberoende.

8. Bevisa att snittet av två underrum till ett vektorrum E också är ett underrum av E. Kan detsamma sägas om unionen? Om inte, ge ett motexempel.

9. Visa att (a) mängden av alla symmetriska matriser av typen n/n, (b) mängden av alla nedåt triangulära matriser av typen n/n, är ett underrum av mängden av alla n/n-matriser. Vad är snittet av dessa två underrum?

10. Beskriv kolonnrummet och nollrummet till matriserna

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{och} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

11. Bestäm echelonformen för matrisen A, då

$$(a)\quad A=\begin{pmatrix}0&1&4&0\\0&2&8&0\end{pmatrix}, \qquad (b)\quad A=\begin{pmatrix}0&0\\1&2\\4&8\\0&0\end{pmatrix}.$$

Bestäm basvariabler och fria variabler i systemet $A\mathbf{x}=\mathbf{0}$. Vilket är A:s nollrum? Vilket villkor bör vara uppfyllt för att systemet $A\mathbf{x}=\mathbf{b}$ skall vara konsistent? Bestäm den allmänna lösningen till detta system då det är konsistent. Vilken rang har A?

12. Är följande delmängder underrum av vektorrummet P av alla polynom:

$$\begin{aligned}&(a)\quad U=\{p\in P \mid p(x)=ax^5,\ a\in\mathbf{R}\}\,;\\ &(b)\quad V=\{p\in P \mid p(x)=2a+x+2x^2,\ a\in\mathbf{R}\}\,;\\ &(c)\quad W=\{p\in P \mid p(-2)=p(1)=0\}\,;\\ &(d)\quad Z=\{p\in P \mid p(2)=p(3)=1\}\,?\end{aligned}$$

13. Visa att mängden M är ett underrum till vektorrummet av alla 2/2-matriser, då

(a) M är mängden av alla matriser X som kommuterar med $A=\begin{pmatrix}1&2\\2&3\end{pmatrix}$;

(b) M är mängden av alla 2/2-matriser av formen $\begin{pmatrix}a&a+b\\0&2b\end{pmatrix}$, $(a,\ b\in\mathbf{R})$;

(c) M är mängden av alla 2/2-matriser av formen $\begin{pmatrix}a&-b\\b&a\end{pmatrix}$, $(a,\ b\in\mathbf{R})$.

(Räkneoperationerna är matrisaddition och multiplikation av en matris med ett reellt tal.)

14. (a) Antag att $\mathbf{a}_1,\ldots,\mathbf{a}_n$ är linjärt oberoende vektorer i ett vektorrum och antag att $\mathbf{a}_{n+1}$ ligger utanför $\operatorname{spn}\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$. Visa att vektorerna $\mathbf{a}_1,\ \ldots,\ \mathbf{a}_{n+1}$ är linjärt oberoende.

(b) Antag att å ena sidan $\{\mathbf{a}_1,\ \ldots,\ \mathbf{a}_m\}$ och å andra sidan $\{\mathbf{b}_1,\ \ldots,\ \mathbf{b}_n\}$ är mängder av linjärt oberoende vektorer i ett vektorrum samt antag att snittet av spannen $\operatorname{spn}\{\mathbf{a}_1,\ldots,\mathbf{a}_m\}$ och $\operatorname{spn}\{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$ är $\{\mathbf{0}\}$. Visa att $\mathbf{a}_1,\ldots,\mathbf{a}_m,\mathbf{b}_1,\ldots,\mathbf{b}_n$ är linjärt oberoende.

15. Antag att $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$ är linjärt oberoende vektorer i ett vektorrum. Är vektorerna

$$\mathbf{a}_1+2\mathbf{a}_2+\mathbf{a}_3,\quad \mathbf{a}_2+2\mathbf{a}_3+\mathbf{a}_4,\quad \mathbf{a}_3+2\mathbf{a}_4+\mathbf{a}_1\quad \mathbf{a}_4+2\mathbf{a}_1+\mathbf{a}_2$$

linjärt beroende? Ange, om de är linjärt beroende, en icke-trivial linjärkombination av dem som är likamed nollvektorn.

16. Bestäm en bas i radrummet, kolonnrummet och nollrummet till matrisen

$$(a)\quad A=\begin{pmatrix}1&-4&9&-7\\-1&2&-4&1\\5&-6&10&7\end{pmatrix}, \qquad (b)\quad A=\begin{pmatrix}-2&-5&8&0&-17\\1&3&-5&1&5\\3&11&-19&7&1\\1&7&-13&5&-3\end{pmatrix}.$$

17. Bestäm dimensionerna för kolonnrummet $R(A)$ och nollrummet $N(A)$ samt ange baser i dessa för alla värden på a, då

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

18. Välj ut en maximal mängd av linjärt oberoende vektorer ur följande vektormängder:

(a) $\{(1\ \ 0\ \ 1\ \ 0), (0\ \ 1\ \ 0\ \ 1), (1\ \ 1\ \ 1\ \ 1), (-1\ \ 0\ \ 2\ \ 1)\}$;
(b) $\{(0\ \ 1\ \ 2\ \ 3), (3\ \ 0\ \ 1\ \ 2), (2\ \ 3\ \ 0\ \ 1), (1\ \ 2\ \ 3\ \ 0)\}$;
(c) $\{(1\ \ -1\ \ 1\ \ -1), (-1\ \ 1\ \ -1\ \ -1), (1\ \ -1\ \ 1\ \ -2), (0\ \ 0\ \ 0\ \ 0)\}$.

19. Ligger $\mathbf{b} = (3\ \ 4\ \ 5)$ i det underrum av $\mathbf{R}^3$ som spänns upp av

$$\mathbf{w}_1 = (1\ \ 1\ \ 0), \quad \mathbf{w}_2 = (2\ \ 2\ \ 1) \quad \text{och} \quad \mathbf{w}_3 = (0\ \ 0\ \ 2)?$$

20. Bevisa att varje delmängd av en mängd av linjärt oberoende vektorer är linjärt oberoende. Gäller det även att en delmängd av en mängd av linjärt beroende vektorer är linjärt beroende?

21. Undersök om följande vektormängder är linjärt beroende eller oberoende

(a) $\{(1\ \ 2\ \ 2), (2\ \ 1\ \ -2), (2\ \ -2\ \ 1)\}$,
(b) $\{(1\ \ 1\ \ 1), (3\ \ 1\ \ 2), (0\ \ 2\ \ 1)\}$,
(c) $\{(1\ \ 1\ \ 1), (3\ \ 1\ \ 2)\}$,
(d) $\{(1\ \ 0\ \ 2), (-2\ \ 0\ \ -4)\}$,
(e) $\{(1\ \ 0\ \ 0), (0\ \ 1\ \ 0), (0\ \ 0\ \ 1), (2\ \ 7\ \ -8)\}$.

Vilka av dessa mängder är en bas i $\mathbf{R}^3$?

22. Visa att $(a\ \ 3)$ och $(2\ \ 1-a)$ är linjärt oberoende vektorer för varje värde på a.

23. Antag att $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_k$ är linjärt oberoende vektorer i ett vektorrum E. Undersök om vektorerna $\mathbf{v}_1 - \mathbf{v}_0$, $\mathbf{v}_2 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0$ är linjärt beroende eller oberoende.

24. Ange en bas och bestäm dimensionen för underrummet av 2/2-matriser definierat i uppgift 13 (a) (använd lösningen av uppgift 5 i kapitel 2), 13 (b) och 13 (c). Vilka är i fallet 13 (b) koordinaterna för

$$\begin{pmatrix} 2 & -1 \\ 0 & -6 \end{pmatrix}$$

i den konstruerade basen?

25. Låt A beteckna en m/n-matris och antag att vi vet att både raderna och kolonnerna i A är linjärt oberoende. Kan vi då säga någonting om typen m/n?
26. Avgör om följande utsagor är sanna eller falska. Ange motexempel i de fall då utsagan är falsk.
    (a) Om $S = \text{spn}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ så är $\dim S = m$.
    (b) Snittet av två underrum kan inte vara den tomma mängden.
    (c) Om $A\mathbf{x} = A\mathbf{y}$ så är $\mathbf{x} = \mathbf{y}$ (A är en matris och $\mathbf{x}$ och $\mathbf{y}$ är kolonnvektorer).
    (d) Radrummet för matrisen A har en entydig bas som kan räknas ut genom att man reducerar A till echelonform.
    (e) Om en kvadratisk matris A har linjärt oberoende kolonner så har även A^2 det.
27. Enligt uppgift 9 (a) är mängden av alla symmetriska 3/3-matriser ett underrum av mängden av alla 3/3-matriser. Bestäm en bas i detta underrum. Vad är underrummets dimension?
28. Visa att om V och W är tredimensionella underrum av $\mathbf{R}^5$ så har V och W en vektor gemensam som är olik nollvektorn.
29. Visa att om $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ är linjärt oberoende så är också $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ linjärt oberoende då vektorerna $\mathbf{w}_i$ är definierade genom $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2$, $\mathbf{w}_2 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$ och $\mathbf{w}_3 = \mathbf{v}_2 + \mathbf{v}_3$.
30. Bevisa att om matriserna A och B är sådana att $AB = 0$, så är kolonnrummet $R(B)$ en delmängd av nollrummet $N(A)$.
31. (a) Visa att polynomen $p_1(x) = 1$, $p_2(x) = 1 + x$ och $p_3(x) = 1 + x + x^2$ bildar en bas i vektorrummet P_2 av alla polynom av högst graden 2. Räkna ut koordinaterna för $p(x) = x^2$ i denna bas.
    (b) Betrakta baserna $B = \{1, x, x^2\}$ och $C = \{p_1, p_2, p_3\}$ i P_2 (se (a)). Antag att polynomet p har koordinaterna 1, 0 och 2 i basen B. Vilka är koordinaterna för p i basen C? Antag att polynomet q har koordinaterna 1, 2 och 5 i basen C. Vilka är koordinaterna för q i basen B?
    (c) Visa att $M = \{1 - x, 1 + x + x^2, 1 - x^2\}$ är en bas i P_2. Vilket polynom har koordinaterna 1, 1 och 1 i basen M.
32. En $2n/2n$-matris M kan delas horisontellt och vertikalt i lika delar så att fyra block A, B, C och D uppstår (varje block är en n/n-matris):
$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$
Om A^{-1} existerar kan man sätta in en nollmatris på C:s plats med en "BO1-operation med matrisblock":
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \to \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$
Vilken matris E skall man multiplicera M med från vänster för att åstadkomma denna operation? Vilken är E:s invers?

33. Använd sådana basoperationer med matrisblock, som beskrivs i uppgift 32, till att invertera matrisen
$$\begin{pmatrix} A & 0 \\ C & D \end{pmatrix},$$
där A och D är inverterbara block.

34. Räkna ut koordinaterna för vektorn $\mathbf{x} = (\,x_1 \quad x_2 \quad x_3 \quad x_4\,)$ i följande bas i $\mathbf{R}^4$:
$$\{(1 \quad 0 \quad 0 \quad 0), (1 \quad 1 \quad 0 \quad 0), (1 \quad 1 \quad 1 \quad 0), (1 \quad 1 \quad 1 \quad 1)\}.$$

35. Bestäm en bas i $\mathbf{R}^2$ sådan att vektorn $(\,3 \quad 7\,)^T$ i denna bas har koordinaterna 2 och 3. Finns det flera baser i $\mathbf{R}^2$, i vilka $(\,3 \quad 7\,)^T$ har de nämnda koordinaterna?

36. *Summan* av två underrum V och W av ett vektorrum E definieras genom
$$V + W = \{\mathbf{x} \in E \mid \mathbf{x} = \mathbf{v} + \mathbf{w},\ \text{där } \mathbf{v} \in V,\ \mathbf{w} \in W\}.$$
Om V är kolonnrummet för en matris A (av typen m/n) och W är kolonnrummet för en matris B (av typen m/p) så är $V + W$ kolonnrummet för den matris $Q = (\,A \quad B\,)$, som man får genom att kombinera A och B till en enda matris. Det gäller då att $\dim(V + W) = r(Q)$. Varför? Bestäm en bas i $V + W$ i det fall att V spänns upp av $\mathbf{v}_1 = (\,1 \quad 1 \quad 0 \quad 0\,)^T$ och $\mathbf{v}_2 = (\,1 \quad 0 \quad 1 \quad 0\,)^T$ samt W spänns upp av $\mathbf{w}_1 = (\,0 \quad 1 \quad 0 \quad 1\,)^T$ och $\mathbf{w}_2 = (\,0 \quad 0 \quad 1 \quad 1\,)^T$. Bestäm också en bas i $V \cap W$ och verifiera i detta speciella fall den allmänna formeln
$$\dim(V + W) + \dim(V \cap W) = \dim V + \dim W. \tag{5}$$

37. Visa att om $R(A) = N(A)$ för en matris A så är denna en kvadratisk matris av en typ n/n, där n är ett jämnt tal.

38. Antag att A är en kvadratisk, inverterbar matris. Bevisa att AB och B har samma nollrum, samma radrum och samma rang.

39. Låt V och W vara underrum av $\mathbf{R}^4$ definierade genom
$$\begin{aligned} V &= \{(\,a \quad b \quad c \quad d\,) \mid b + c + d = 0\}, \\ W &= \{(\,a \quad b \quad c \quad d\,) \mid a + b = 0,\ c = 2d\}. \end{aligned}$$
Ange en bas för de fyra underrummen V, W, $V \cap W$ och $V + W$ samt verifiera att formel (5) gäller ($V + W$ definieras i uppgift 36).

40. Visa att vektorerna $\mathbf{v} = (\,1 \quad 0 \quad 1 \quad 0\,)$ och $\mathbf{w} = (\,0 \quad 1 \quad 0 \quad 1\,)$ tillhör underrummet U av $\mathbf{R}^4$ definierat genom $U = \{(\,x \quad y \quad z \quad u\,) \mid x + y = z + u\}$. Är mängden $\{\mathbf{v}, \mathbf{w}\}$ en bas i U? Om inte, komplettera den med ett nödvändigt antal vektorer till en bas i U.

41. Visa att matriserna
$$\begin{pmatrix} 2 & 3 & 4 \\ 0 & 1 & 1 \\ 2 & -1 & 0 \end{pmatrix} \quad \text{och} \quad \begin{pmatrix} 4 & 2 & 4 \\ 2 & 0 & 1 \\ 2 & 7 & 8 \end{pmatrix}$$

har samma radrum och ange en bas i detta.

42. Har följande matris någon högerinvers (motivera!):

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & -1 & 3 & 2 \\ 1 & 5 & -1 & 4 \end{pmatrix}?$$

43. Konstruera alla möjliga vänster- och högerinverser till matrisen $A = (1 \quad 1 \quad 0)$.

44. Visa att matriserna

$$A = \begin{pmatrix} 2 & -2 \\ 4 & -4 \\ 0 & 0 \end{pmatrix} \quad \text{och} \quad B = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 3 & 6 \end{pmatrix}$$

har rangen 1 och skriv dem i formen $A = \mathbf{u}\mathbf{v}^T$ och $B = \mathbf{w}\mathbf{z}^T$. Verifiera att deras produkt AB är en multipel av matrisen $\mathbf{u}\mathbf{z}^T$ och att denna multipel är $\mathbf{v}^T\mathbf{w}$.

5. Euklidiska vektorrum

Avstånd och vinklar kan inte definieras enbart med hjälp av räkneoperationerna i ett vektorrum E, dvs. med hjälp av addition och multiplikation med skalär. Vi skall nu lägga till mera struktur i E genom att införa en ny räkneoperation, den skalära produkten. Genom denna nya produkt, som bör underlyda vissa räkneregler (axiom), kan både avstånds- och vinkelbegreppet definieras.

Den skalära produkten införs genom axiom eftersom det finns oändligt många sätt att definiera en sådan produkt. Genom detta kommer våra satser och bevis att gälla för varje skalär produkt som uppfyller axiomen.

Definition 5.1. Ett vektorrum E kallas ett *euklidiskt* vektorrum om det är försett med en så kallad *skalär produkt*: Till varje par av vektorer $\mathbf{x}$, $\mathbf{y} \in E$ hör en entydig skalär $(\mathbf{x}, \mathbf{y})$, kallad den skalära produkten av $\mathbf{x}$ och $\mathbf{y}$, sådan att följande axiom gäller:

IV (*a*) $(\mathbf{x}, \mathbf{x}) \geq 0$ för varje $\mathbf{x} \in E$; $\quad (\mathbf{x}, \mathbf{x}) = 0$ bara om $\mathbf{x} = \mathbf{0}$;
(*b*) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$ för varje $\mathbf{x}$, $\mathbf{y} \in E$;
(*c*) $(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$, för varje $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in E$;
(*d*) $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$, för varje $\mathbf{x}$, $\mathbf{y} \in E$.

Vi ser genast att $(\mathbf{0}, \mathbf{y}) = 0$ för varje $y \in E$, ty $(\mathbf{0}, \mathbf{y}) = (0\mathbf{x}, \mathbf{y}) = 0(\mathbf{x}, \mathbf{y}) = 0$. Den andra delen av axiom IV (*a*) kunde alltså lika gärna skrivas: $(\mathbf{x}, \mathbf{x}) = 0$ om och endast om $\mathbf{x} = \mathbf{0}$.

Anmärkning. I litteraturen förekommer också beteckningar som $\langle \mathbf{x}, \mathbf{y} \rangle$, $\langle \mathbf{x} \mid \mathbf{y} \rangle$ och $(\mathbf{x} \mid \mathbf{y})$ för den skalära produkten av $\mathbf{x}$ och $\mathbf{y}$.

Exempel 5.1. Vi kan göra $\mathbf{R}^n$ till ett euklidiskt vektorrum genom att definiera en skalär produkt genom

$$(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i \,.$$

Denna brukar kallas den *naturliga skalära produkten* och kan skrivas som en matrisprodukt

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$$

ifall vi uppfattar faktorerna som kolonnvektorer (t.ex. $\mathbf{x} = (\, x_1 \;\; \ldots \;\; x_n \,)^T$). Att axiomen IV (*a*) – (*d*) gäller, följer ur räknereglerna för matrisprodukter.

Exempel 5.2. I $\mathbf{R}^3$ kan vi införa en skalär produkt t.ex. genom att sätta

$$(\mathbf{x}, \mathbf{y}) = 3x_1 y_1 + 2x_2 y_2 + 11 x_3 y_3 = \mathbf{x}^T \begin{pmatrix} 3 & & \\ & 2 & \\ & & 11 \end{pmatrix} \mathbf{y} \,.$$

Observera här att diagonalelementen i diagonalmatrisen måste vara positiva för att axiom IV (a) skall gälla. Byter man alltså ut talet 11 mot -11 så duger det uppkomna uttrycket inte som en skalär produkt.

Exempel 5.3. Låt oss noggrant verifiera att om vi i $\mathbf{R}^2$ sätter

$$(\mathbf{x},\mathbf{y}) = \mathbf{x}^T A\mathbf{y} = \mathbf{x}^T \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \mathbf{y}$$

så har vi definierat en skalär produkt: Med hjälp av kvadratkomplettering får vi

$$\begin{aligned}(\mathbf{x},\mathbf{x}) &= 3x_1^2 + 4x_1x_2 + 3x_2^2 = 3(x_1^2 + \frac{4}{3}x_1x_2) + 3x_2^2 \\ &= 3\left(x_1^2 + \frac{4}{3}x_1x_2 + \frac{4}{9}x_2^2\right) + \left(3 - \frac{4}{3}\right)x_2^2 = 3\left(x_1 + \frac{2}{3}x_2\right)^2 + \frac{5}{9}x_2^2 \geq 0\end{aligned}.$$

Ur detta följer att $(\mathbf{x},\mathbf{x}) = 0$ om och endast om både $x_1 + (2/3)x_2 = 0$ och $x_2 = 0$, dvs. om och endast om $x_1 = x_2 = 0$. Alltså gäller axiom IV (a). Vid transponering av en 1/1-matris (= skalär) förblir matrisen oförändrad. Detta tillsammans med det faktum att A är symmetrisk ger att IV (b) gäller:

$$(\mathbf{x},\mathbf{y}) = \mathbf{x}^T A\mathbf{y} = \left(\mathbf{x}^T A\mathbf{y}\right)^T = \mathbf{y}^T A^T \mathbf{x}^{TT} = \mathbf{y}^T A\mathbf{x} = (\mathbf{y},\mathbf{x}).$$

Välkända räkneregler för matriser ger giltigheten hos de två sista axiomen:

$$\begin{aligned}(\mathbf{x}+\mathbf{y},\mathbf{z}) &= (\mathbf{x}+\mathbf{y})^T A\mathbf{z} = (\mathbf{x}^T + \mathbf{y}^T)A\mathbf{z} = \mathbf{x}^T A\mathbf{z} + \mathbf{y}^T A\mathbf{z} = (\mathbf{x},\mathbf{y}) + (\mathbf{y},\mathbf{z}) \\ (\lambda\mathbf{x},\mathbf{y}) &= (\lambda\mathbf{x})^T A\mathbf{y} = \lambda\mathbf{x}^T A\mathbf{y} = \lambda(\mathbf{x},\mathbf{y}).\end{aligned}$$

Observera att det är likgiltigt vilka matriselement A innehåller bara regel IV (a) gäller och A är symmetrisk.

Exempel 5.4. Låt

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2 + a_3x^3 \quad \text{och} \\ q(x) &= b_0 + b_1x + b_2x^2 + b_3x^3 \end{aligned}$$

vara två godtyckliga polynom i vektorrummet P_3 av polynom av högst graden 3. Vi kan då införa en skalär produkt i P_3 på många olika sätt, t.ex. genom att sätta

$$(p,q) = a_0b_0 + 3a_1b_1 + 5a_2b_2 + 7a_3b_3\,,$$

eller genom att sätta

$$(p,q) = \int_{-2}^{9} x^4 p(x)q(x)\,dx\,.$$

Man bevisar detta genom att i tur och ordning verifiera axiomen IV $(a) - (d)$ med hjälp kända räkneregler (för t.ex. integraler).

Avstånd och vinklar

Definition 5.2. Med *normen* (eller *längden*) av en vektor $\mathbf{x}$ i ett euklidiskt vektorrum E avses talet

$$\|\mathbf{x}\| = +\sqrt{(\mathbf{x}, \mathbf{x})}\,.$$

Avståndet mellan två vektorer $\mathbf{x}$ och $\mathbf{y}$ i E är talet

$$\|\mathbf{x} - \mathbf{y}\|\,.$$

En vektor $\mathbf{x}$ sägs vara en *enhetsvektor* om $\|\mathbf{x}\| = 1$.

Enligt dessa definitioner är normen av en vektor $\mathbf{x}$ detsamma som avståndet från $\mathbf{0}$ till $\mathbf{x}$. Observera att norm och avstånd är beroende av den skalära produkten i vektorrummet. Olika skalära produkter ger i allmänhet upphov till olika talvärden för norm och avstånd:

Exempel 5.5. I $\mathbf{R}^2$ svarar den skalära produkten

$$\mathbf{x}^T \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix} \mathbf{y} \quad \text{mot normen} \quad \|\mathbf{x}\| = \sqrt{5x_1^2 + 2x_2^2}\,.$$

Mot den skalära produkten

$$\mathbf{x}^T \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} \mathbf{y} \quad \text{svarar normen} \quad \|\mathbf{x}\| = \sqrt{2x_1^2 + 4x_1x_2 + 3x_2^2}\,.$$

Eftersom

$$\|\lambda\mathbf{x}\|^2 = (\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda(\mathbf{x}, \lambda\mathbf{x}) = \lambda(\lambda\mathbf{x}, \mathbf{x}) = \lambda^2(\mathbf{x}, \mathbf{x}) = \lambda^2\|\mathbf{x}\|^2$$

så gäller räkneregeln

$$\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$$

för varje norm. För att kunna härleda fler räkneregler, behöver vi Schwarz olikhet:

Sats 5.1. (Schwarz olikhet) *I ett euklidiskt vektorrum E gäller för varje skalär produkt och motsvarande norm olikheterna*

$$(\mathbf{x}, \mathbf{y}) \leq |(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

för varje $\mathbf{x} \in E$ och $\mathbf{y} \in E$.

Bevis. Vi kan allra först konstatera att alla tre leden blir noll om $\mathbf{x} = \mathbf{0}$ eller $\mathbf{y} = \mathbf{0}$. Påståendet gäller alltså i sådana fall. Låt oss därför i fortsättningen av beviset anta att $\mathbf{x} \neq \mathbf{0}$ och $\mathbf{y} \neq \mathbf{0}$.

Vi skall minimera kvadraten av avståndet,

$$f(t) = \|t\mathbf{x} - \mathbf{y}\|^2\,,$$

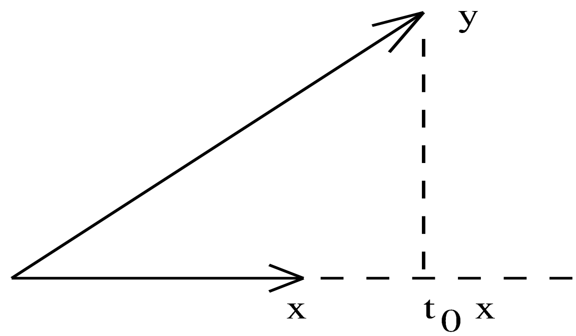

fig. 1

mellan $t\mathbf{x}$ och $\mathbf{y}$ genom att variera t.

Eftersom en kvadrat inte kan vara negativ, är $f(t_0) \geq 0$ för det värde t_0 som ger minimum. Denna olikhet är just Schwarz olikhet. Vi kan bestämma t_0 t.ex. genom kvadratkomplettering:

$$\begin{aligned}
f(t) &= \|t\mathbf{x}-\mathbf{y}\|^2 = (t\mathbf{x}-\mathbf{y}, t\mathbf{x}-\mathbf{y}) \\
&= (t\mathbf{x}, t\mathbf{x}) - (t\mathbf{x}, \mathbf{y}) - (\mathbf{y}, t\mathbf{x}) + (\mathbf{y}, \mathbf{y}) \\
&= t^2\|\mathbf{x}\|^2 - 2t(\mathbf{x}, \mathbf{y}) + \|\mathbf{y}\|^2 \\
&= \|\mathbf{x}\|^2 \left(t^2 - 2t\frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|^2} + \frac{(\mathbf{x}, \mathbf{y})^2}{\|\mathbf{x}\|^4} \right) + \|\mathbf{y}\|^2 - \frac{(\mathbf{x}, \mathbf{y})^2}{\|\mathbf{x}\|^2} \\
&= \|\mathbf{x}\|^2 \left(t - \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|^2} \right)^2 + \|\mathbf{y}\|^2 - \frac{(\mathbf{x}, \mathbf{y})^2}{\|\mathbf{x}\|^2} .
\end{aligned}$$

Det minsta värdet för $f(t)$ antas då

$$t = t_0 = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|^2} .$$

Detta ger oss olikheten

$$0 \leq f(t_0) = \|\mathbf{y}\|^2 - \frac{(\mathbf{x}, \mathbf{y})^2}{\|\mathbf{x}\|^2} ,$$

vilket efter omformning och kvadratrotsutdragning ger $|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|\|\mathbf{y}\|$. Den första olikheten i formeln i satsen är självklar, eftersom $c \leq |c|$ för varje skalär c. $\diamondsuit$

Exempel 5.6. Mot den skalära produkten

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} \mathbf{y}$$

(se exempel 5.5) svarar en Schwarz-olikhet med utseendet

$$|2x_1y_1 + 2x_1y_2 + 2x_2y_1 + 3x_2y_2| \leq \sqrt{2x_1^2 + 4x_1x_2 + 3x_2^2}\sqrt{2y_1^2 + 4y_1y_2 + 3y_2^2} .$$

Med hjälp av Schwarz olikhet fås den s.k. *triangelolikheten*, en viktig räkneregel för normen,

$$\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|,$$

ty

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y},\, \mathbf{x}+\mathbf{y}) = (\mathbf{x},\, \mathbf{x}) + 2(\mathbf{x},\, \mathbf{y}) + (\mathbf{y},\, \mathbf{y}) \\ &= \|\mathbf{x}\|^2 + 2(\mathbf{x},\, \mathbf{y}) + \|\mathbf{y}\|^2 \le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\cdot\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 .\end{aligned}$$

Det är naturligt att säga att det kortaste avståndet från $\mathbf{y}$ till $\operatorname{spn}\{\mathbf{x}\}$ i fig. 1 är det "vinkelräta avståndet". Genom detta fastslås både begreppet rät vinkel och vinkelmåttet ϕ för vinkeln mellan vektorerna $\mathbf{x}$ och $\mathbf{y}$, ty ur den "rätvinkliga" triangeln i fig. 1 fås

$$\cos\varphi = \pm\frac{\|t_0\mathbf{x}\|}{\|\mathbf{y}\|} = \pm|t_0|\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|} = t_0\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|},$$

där $t_0 = (\mathbf{x},\, \mathbf{y})/\|\mathbf{x}\|^2$ enligt beviset för Schwarz olikhet. Den naturliga definitionen för vinklar är alltså:

Definition 5.3. *Vinkeln* φ mellan två vektorer $\mathbf{x}$ och $\mathbf{y}$, $\mathbf{x} \neq \mathbf{0}$ och $\mathbf{y} \neq \mathbf{0}$, fås ur formeln

$$\cos\varphi = \frac{(\mathbf{x},\, \mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

Vinkeln blir entydig om vi kommer överens om att alla vinklar skall ha ett värde i intervallet $[0,\, \pi]$. Definitionslikheten kan också skrivas

$$(\mathbf{x},\, \mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\varphi.$$

Anmärkning. Lägg märke att ett vinkelmått existerar så snart $\mathbf{x} \neq \mathbf{0}$ och $\mathbf{y} \neq \mathbf{0}$, eftersom högerledet i definitionsformeln då ligger i intervallet $[-1, 1]$ enligt Schwarz olikhet. Om någondera vektorn är nollvektorn, existerar ingen vinkel. Lägg också märke till att vinkelmåttet (precis som avståndsmåttet) helt är beroende av vilken skalär produkt som används.

Exempel 5.7. Antag att $\mathbf{R}^3$ är försett med den naturliga skalära produkten. För vektorerna $\mathbf{x} = (1 \quad 1 \quad 4)^T$ och $\mathbf{y} = (4 \quad 1 \quad 1)^T$ är $\|\mathbf{x}\| = \|\mathbf{y}\| = \sqrt{18}$ och $(\mathbf{x},\, \mathbf{y}) = 9$. För vinkeln φ mellan dem gäller därför

$$\cos\varphi = \frac{9}{\sqrt{18}\sqrt{18}} = \frac{1}{2},$$

varför φ är $\pi/3$, dvs. 60°. Om däremot $\mathbf{R}^3$ är försett med t.ex. skalärprodukten i exempel 5.2 så fås ett annat värde på vinkeln mellan $\mathbf{x}$ och $\mathbf{y}$.

Definition 5.4. Två vektorer $\mathbf{x}$ och $\mathbf{y}$ i ett euklidiskt vektorrum är *ortogonala* om $(\mathbf{x},\, \mathbf{y}) = 0$. Detta betecknas $\mathbf{x} \perp \mathbf{y}$.

Om $\mathbf{x} \neq \mathbf{0}$ och $\mathbf{y} \neq \mathbf{0}$ är dessa två vektorer ortogonala om och endast om vinkeln mellan dem är rät, dvs. likamed $\pi/2 = 90^\circ$. Observera att nollvektorn enligt vår definition är ortogonal mot varje vektor i vektorrummet, eftersom $(\mathbf{0},\, \mathbf{x}) = 0$.

För den naturliga skalära produkten i $\mathbf{R}^n$ gäller att $\mathbf{x} = (\,x_1 \quad \ldots \quad x_n\,)$ och $\mathbf{y} = (\,y_1 \quad \ldots \quad y_n\,)$ är ortogonala om

$$\mathbf{x}^T\mathbf{y} = x_1y_1 + \cdots + x_ny_n = 0\,.$$

Definition 5.5. Vi säger att ett antal vektorer $\mathbf{x}_1, \ldots, \mathbf{x}_k$ i ett euklidiskt vektorrum är *ortogonala* om de är parvis ortogonala, dvs. om $\mathbf{x}_i \perp \mathbf{x}_j$ för $i \neq j$, och *ortonormala* om det dessutom gäller att varje $\mathbf{x}_i$ är en enhetsvektor, dvs. om $\|\mathbf{x}_i\| = 1$. Ortonormala vektorer sägs bilda ett *ortonormerat system* (förkortat: *ON-system*).

Exempel 5.8. Låt $\mathbf{R}^2$ vara försett med den naturliga skalära produkten. Om ON-systemet $\{(\,1 \quad 0\,)^T, (\,0 \quad 1\,)^T\}$ av naturliga enhetsvektorer vrids en vinkel θ, fås ett nytt ON-system $\{\mathbf{v}_1, \mathbf{v}_2\}$.

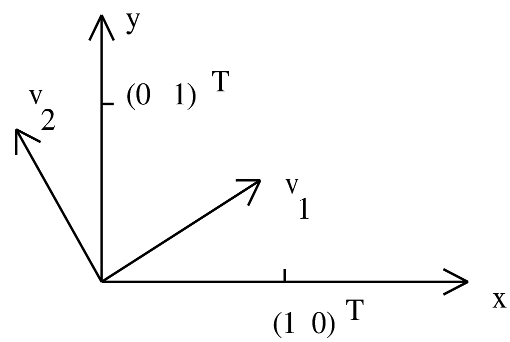

fig. 2

Sätt $\mathbf{v}_1 = (\,x_1 \quad y_1\,)^T$ och $\mathbf{v}_2 = (\,x_2 \quad y_2\,)^T$. Då är

$$\begin{aligned} x_1 &= \cos\theta, \quad y_1 = \sin\theta \\ x_2 &= -\sin\theta, \quad y_2 = \cos\theta\,. \end{aligned}$$

Det nya ON-systemet är alltså $\{(\,\cos\theta \quad \sin\theta\,)^T, (\,-\sin\theta \quad \cos\theta\,)^T\}$.

Exempel 5.9. Polynomen $1/\sqrt{2}$, $\sqrt{3/2}\,x$ bildar ett ON-system i vektorrummet P av alla polynom om den skalära produkten är

$$(p,q) = \int_{-1}^{1} p(x)q(x)\,dx\,.$$

Sats 5.2. *Om vektorerna $\mathbf{v}_1, \ldots, \mathbf{v}_k$ är ortogonala i ett euklidiskt vektorrum och $\mathbf{v}_i \neq \mathbf{0}$ för varje $i = 1, \ldots, k$ så är $\mathbf{v}_1, \ldots, \mathbf{v}_k$ linjärt oberoende.*

Bevis. Antag att $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$. Genom att bilda den skalära produkten av bägge leden med $\mathbf{v}_i$, fås

$$c_1(\mathbf{v}_i, \mathbf{v}_1) + \cdots + c_i(\mathbf{v}_i, \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_i, \mathbf{v}_k) = 0\,.$$

I vänstra ledet är alla termer noll utom den i:te. Således är $c_i\|\mathbf{v}_i\|^2 = 0$ och följaktligen $c_i = 0$, eftersom $\mathbf{x}_i \neq 0$. Men detta gäller för varje $i \in \{1, \ldots, k\}$. Alltså är $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ linjärt oberoende. $\diamondsuit$

Exempel 5.10. Två linjärt oberoende vektorer i $\mathbf{R}^2$ behöver naturligtvis inte automatiskt vara ortogonala i den naturliga skalära produkten. Däremot existerar det för två givna linjärt oberoende vektorer $\mathbf{a}$ och $\mathbf{b}$ alltid en annan skalär produkt, i vilken $\mathbf{a}$ och $\mathbf{b}$ blir ortonormala: Låt ξ_1, ξ_2 respektive η_1, η_2 beteckna koordinaterna i basen $\{\mathbf{a}, \mathbf{b}\}$ för godtyckliga vektorer $\mathbf{x}$ och $\mathbf{y}$ och definiera en skalär produkt genom

$$(\mathbf{x}, \mathbf{y}) = \xi_1\eta_1 + \xi_2\eta_2 .$$

Eftersom $\mathbf{a}$ och $\mathbf{b}$ har koordinaterna $1, 0$ respektive $0, 1$, är uppenbarligen $(\mathbf{a}, \mathbf{b}) = 0$. För de nya normerna gäller $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$.

Geometriska objekt i $\mathbf{R}^n$

I fortsättningen *använder vi i* $\mathbf{R}^n$ *alltid den naturliga skalära produkten, ifall ingenting annat uttryckligen sägs.* Detta gäller genom resten av kompendiet.

(a) En *rät linje* L är entydigt bestämd genom att man anger en fast punkt $\mathbf{x}_0$ på L och en riktning $\mathbf{v} \neq 0$ för L. En punkt $\mathbf{x}$ ligger på L, dvs. $\mathbf{x} \in L$, om och endast om

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$$

för något $t \in \mathbf{R}$. Detta är linjens ekvation i parameterform med t som parameter.

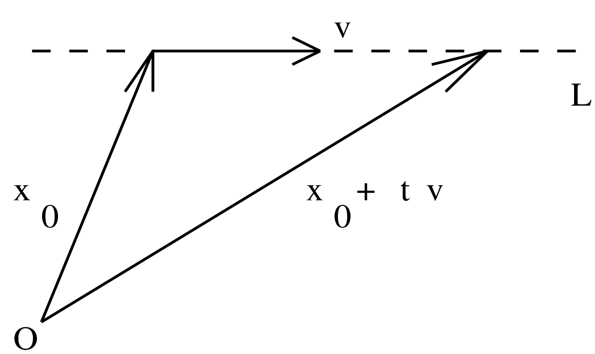

fig. 3

Om L ligger t.ex. i $\mathbf{R}^3$ och om $\mathbf{x} = (x \quad y \quad z)^T$, $\mathbf{x}_0 = (x_0 \quad y_0 \quad z_0)^T$ och $\mathbf{v} = (v_1 \quad v_2 \quad v_3)^T$, kan ekvationen skrivas i komponentform

$$\begin{aligned} x &= x_0 + tv_1 \\ y &= y_0 + tv_2 \\ z &= z_0 + tv_3 . \end{aligned}$$

Om linjen L är definierad av att den går genom $\mathbf{x}_0$ och $\mathbf{x}_1$ ($\neq \mathbf{x}_0$) så kan man som riktningvektor välja $\mathbf{v} = \mathbf{x}_1 - \mathbf{x}_0$.

(*b*) Ett *plan* P i $\mathbf{R}^3$ kan definieras av en punkt $\mathbf{x}_0$ som ligger i P och två riktningar $\mathbf{u}$ och $\mathbf{v}$. Planets ekvation i parameterform blir $\mathbf{x} = \mathbf{x}_0 + s\mathbf{u} + t\mathbf{v}$. Alternativt kan P:s läge anges genom $\mathbf{x}_0$ och en normalvektor $\mathbf{n} = (\,l \quad m \quad n\,)^T$ till P.

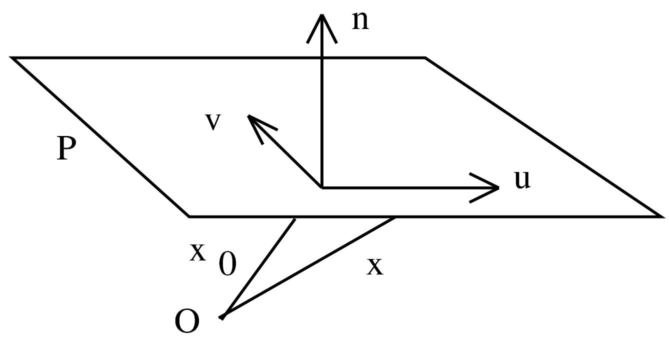

fig. 4

En punkt $\mathbf{x} = (\,x \quad y \quad z\,)^T$ ligger i P om och endst om vektorn $\mathbf{x} - \mathbf{x}_0$ är ortogonal mot $\mathbf{n}$, dvs. om och endast om

$$\mathbf{n}^T(\mathbf{x} - \mathbf{x}_0) = 0\,. \tag{1}$$

Detta är ekvationen för planet P. Fullständigt utskriven blir ekvationen

$$l(x - x_0) + m(y - y_0) + n(z - z_0) = 0\,,$$

dvs. den är av formen $Ax + By + Cz + D = 0$. Varje ekvation av denna form, där $\mathbf{N} = (\,A \quad B \quad C\,)^T \neq \mathbf{0}$, representerar uppenbarligen ett plan med normalvektorn $\mathbf{N}$.

Om dimensionen för $\mathbf{R}^n$ är större än 3 säger man att (1) är ekvationen för ett *hyperplan* i $\mathbf{R}^n$.

Exempel 5.11. I $\mathbf{R}^4$ är $U = \{(\,x_1 \quad 0 \quad x_3 \quad x_4\,)^T \,|x_1, x_3, x_4 \in \mathbf{R}\}$ ett hyperplan genom origo, i detta fall ett underrum med dimensionen 3. Det har ekvationen $x_2 = 0$ och normalvektorn $(\,0 \quad 1 \quad 0 \quad 0\,)^T$. Ekvationen

$$2(x_1 - 1) - (x_2 - 2) + 3(x_3 + 1) + x_4 = 0$$

representerar ett hyperplan genom punkten $(\,1 \quad 2 \quad -1 \quad 0\,)^T$ med normalvektorn $(\,2 \quad -1 \quad 3 \quad 1\,)^T$. Ekvationen kan också skrivas $2x_1 - x_2 + 3x_3 + x_4 + 3 = 0$. Om denna variant av ekvationen är given, hittar man lätt en punkt på hyperplanet genom att ge tre variabler (fria variabler) godtyckliga värden och sedan räkna ut den fjärde. En sådan punkt är t.ex. $(\,-3/2 \quad 0 \quad 0 \quad 0\,)^T$.

Det är lätt att se (och har framgått genom exempel) att **en rät linje respektive ett hyperplan i $\mathbf{R}^n$ är ett underrum om och endast om den/det går genom origo**.

Ortogonala underrum

Definition 5.6. Två underrum V och W av ett euklidiskt vektorrum sägs vara *ortogonala* om $\mathbf{v} \perp \mathbf{w}$ för varje $\mathbf{v} \in V$ och $\mathbf{w} \in W$. För ortogonaliteten använder vi beteckningen $V \perp W$

Exempel 5.12. I $\mathbf{R}^3$ är xy-planet och z-axeln ortogonala. I $\mathbf{R}^4$ är underrummen

$$V = \text{spn}\left\{(1 \quad 1 \quad 0 \quad 0)^T, (0 \quad 0 \quad 1 \quad 1)^T\right\}$$
$$W = \text{spn}\left\{(1 \quad -1 \quad 0 \quad 0)^T, (0 \quad 0 \quad 1 \quad -1)^T\right\}$$

ortogonala, ty de två vektorer som spänner upp V är bägge ortogonala mot de två vektorer som spänner upp W.

Sats 5.3. *För varje m/n-matris A gäller*

$$\begin{array}{lll} (a) & N(A) \perp R(A^T) & \text{i } \mathbf{R}^n \\ (b) & N(A^T) \perp R(A) & \text{i } \mathbf{R}^m . \end{array}$$

Bevis. Relation (b) följer ur relation (a) genom att A ersätts med A^T. För att bevisa relation (a) väljer vi godtyckliga vektorer $\mathbf{x} \in N(A)$ och $\mathbf{y} \in R(A^T)$. Då gäller att $A\mathbf{x} = \mathbf{0}$ och att $\mathbf{y} = A^T\mathbf{u}$ för något $\mathbf{u} \in \mathbf{R}^m$. För den skalära produkten fås nu

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y} = \mathbf{x}^T A^T \mathbf{u} = (A\mathbf{x})^T\mathbf{u} = \mathbf{0}^T\mathbf{u} = \mathbf{0}.$$

Underrummen $N(A)$ och $R(A^T)$ är alltså ortogonala. $\diamondsuit$

Definition 5.7. Antag att V är ett underrum av ett euklidiskt vektorrum E. Mängden

$$V^\perp = \{\mathbf{y} \in E \mid \mathbf{y} \perp \mathbf{x} \text{ för varje } \mathbf{x} \in V\}$$

av alla vektorer $\mathbf{y}$ som är ortogonala mot V kallas det *ortogonala komplementet* till V.

Det är lätt att se att $V^\perp$ är ett underrum av E. Direkt ur definitionen följer också att ett underrum V och dess ortogonala komplement $V^\perp$ alltid är ortogonala underrum:

$$V \perp V^\perp .$$

Exempel 5.13. Observera att två underrum kan vara ortogonala utan att vara varandras ortogonala komplement: Om V är x-axeln och W är y-axeln i $\mathbf{R}^3$ så är W inte det det ortogonala komplementet till V, ty $V^\perp$ omfattar uppenbarligen hela yz-planet.

Vi kan nu precisera relation (a) i Sats 5.3:

Sats 5.4. *Antag att A är en m/n-matris. Då gäller relationerna:*

$$(a) \quad R(A^T)^\perp = N(A)\,,$$
$$(b) \quad R(A^T) = N(A)^\perp\,.$$

Bevis. (a) Låt $\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_m$ beteckna raderna i A. Då är

$$\begin{aligned} \mathbf{x} \in N(A) \quad &\Longleftrightarrow \quad A\mathbf{x} = \mathbf{0} \\ &\Longleftrightarrow \quad \bar{\mathbf{a}}_i\mathbf{x} = 0 \text{ för } i = 1, \ldots, m \\ &\Longleftrightarrow \quad \left(\sum_{i=1}^m c_i\bar{\mathbf{a}}_i\right)\mathbf{x} = 0 \text{ för varje } \mathbf{y} = \sum_{i=1}^m c_i\bar{\mathbf{a}}_i \in R(A^T) \\ &\Longleftrightarrow \quad \mathbf{x} \in R(A^T)^\perp\,. \end{aligned}$$

Således är $N(A) = R(A^T)^\perp$.

(b) Eftersom $N(A) \perp R(A^T)$ enligt Sats 5.3, är $R(A^T) \subseteq N(A)^\perp$. Vi skall visa att denna inklusion i själva verket är en likhet. För detta antar vi som en **antites** att det finns en radvektor $\bar{\mathbf{z}} \in N(A)^\perp$ sådan att $\bar{\mathbf{z}} \notin R(A^T)$. Sätt

$$A_1 = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_m \\ \bar{\mathbf{z}} \end{pmatrix}.$$

Då är det klart att $N(A) = N(A_1)$, eftersom vektorn $\bar{\mathbf{z}}$ valts ortogonal mot $N(A)$. Men ur $\bar{\mathbf{z}} \notin R(A^T)$ följer att $\dim R(A_1^T) = \dim R(A^T) + 1$ och vidare enligt Sats 4.13 att $\dim R(A_1) = \dim R(A) + 1$. Därmed är, enligt Sats 4.14,

$$n = \dim R(A_1) + \dim N(A_1) = \dim R(A) + \dim N(A) + 1 = n + 1\,,$$

eftersom både A och A_1 har n kolonner. Men detta är en motsägelse. Slutsatsen blir att $R(A^T) = N(A)^\perp$. $\diamondsuit$

Sats 5.4 kan naturligtvis tillämpas också på matrisen A^T. Man får då en precisering av relation (b) i Sats 5.3:

$$R(A)^\perp = N(A^T)\,, \qquad R(A) = N(A^T)^\perp\,.$$

Vi kan nu bevisa:

Sats 5.5. *Antag att V och W är underrum av $\mathbf{R}^n$.*

(i) Följande påståenden är ekvivalenta:

$$(1) \quad W = V^\perp\,,$$
$$(2) \quad V = W^\perp\,,$$
$$(3) \quad V \perp W \text{ och } \dim V + \dim W = n\,.$$

(ii) Då (1), (2) eller (3) gäller, har varje $\mathbf{x} \in \mathbf{R}^n$ en entydig uppdelning

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2\,, \text{ där } \mathbf{x}_1 \in V\,,\ \mathbf{x}_2 \in W\,.$$

Vektorn $\mathbf{x}_1$ kallas härvid den ortogonala projektionen *av $\mathbf{x}$ på V och $\mathbf{x}_2$ kallas den* ortogonala projektionen *av $\mathbf{x}$ på W.*

Bevis. (*i*) Antag att (1) gäller. Låt $\{\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_k\}$ vara en bas i V och sätt

$$A = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_k \end{pmatrix}.$$

Då är $R(A^T) = V$ och $N(A) = W$. Enligt Sats 5.4 är nu $N(A)^\perp = R(A^T)$, så att $W^\perp = V$, dvs. (2) gäller. På samma sätt följer (1) ur (2).

Om något av de ekvivalenta påståendena (1) eller (2) gäller, är $V \perp W$ och för matrisen A ovan fås att

$$\dim V + \dim W = \dim R(A^T) + \dim N(A) = \dim R(A) + \dim N(A) = n$$

dvs. (3) gäller. Antag omvänt att (3) gäller för underrummen V och W och sätt $W_1 = V^\perp$. Då är $W \subseteq W_1$. Enligt vad vi just visat är $\dim V + \dim W_1 = n$. Med hjälp av (3) fås nu att $\dim W = \dim W_1$. Detta tillsammans med inklusionen $W \subseteq W_1$ ger likheten $W = W_1 = V^\perp$. Alltså gäller (1).

(*ii*) Tag en bas $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ i V och en bas $\{\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ i W. Då består unionen av dessa baser av linjärt oberoende vektorer i $\mathbf{R}^n$, ty om $c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{0}$ så är

$$c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k = -(c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n),$$

där vänsterledet tillhör V medan högerledet tillhör W. Men eftersom V och W bara har vektorn $\mathbf{0}$ gemensam (de är ju ortogonala) så är bägge leden $\mathbf{0}$. Detta ger att att $c_1 = \cdots = c_k = 0$ och att $c_{k+1} = \cdots = c_n = 0$. Vektorerna b_i $(i = 1, \ldots, n)$ bildar därför enligt Sats 4.7 en bas i $\mathbf{R}^n$. Varje vektor $\mathbf{x} \in \mathbf{R}^n$ är därmed en linjärkombination

$$\mathbf{x} = (d_1\mathbf{b}_1 + \cdots + d_k\mathbf{b}_k) + (d_{k+1}\mathbf{b}_{k+1} + \cdots + d_n\mathbf{b}_n) = \mathbf{x}_1 + \mathbf{x}_2,$$

där $\mathbf{x}_1 \in V$ och $\mathbf{x}_2 \in W$.

Entydigheten hos denna uppdelning fås på följande sätt: Antag att $\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2$, där $\mathbf{y}_1 \in V$ och $\mathbf{y}_2 \in W$, är en godtycklig uppdelning av $\mathbf{x}$. Då är ju $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{y}_1 + \mathbf{y}_2$, dvs.

$$\mathbf{x}_1 - \mathbf{y}_1 = \mathbf{y}_2 - \mathbf{x}_2,$$

där vänsterledet ligger i V och högerledet i W. Ortogonaliteten hos V och W medför att bägge leden är $\mathbf{0}$, så att $\mathbf{x}_1 = \mathbf{y}_1$ och $\mathbf{x}_2 = \mathbf{y}_2$. Varje uppdelning är alltså identisk med den som vi först presenterade. $\diamondsuit$

Anmärkning. Med samma beteckningar som i Sats 5.5 fås genom insättning av $W = V^\perp$ i likheten $V = W^\perp$ att

$$V = V^{\perp\perp}.$$

Denna relation gäller för vilket underrum V av $\mathbf{R}^n$ som helst.

Exempel 5.15. I detta exempel skall vi visa hur man bestämmer en bas i ett ortogonalt komplement. Antag att

$$V = \operatorname{spn}\left\{(1 \quad -2 \quad 3)^T, (2 \quad 3 \quad -1)^T\right\}.$$

Då består $V^\perp$ av alla $\mathbf{x} \in \mathbf{R}^3$, sådana att $\mathbf{a}_1^T\mathbf{x} = 0$ och $\mathbf{a}_2^T\mathbf{x} = 0$. Detta betyder att $V^\perp$ är nollrummet till matrisen med raderna $\mathbf{a}_1^T$ och $\mathbf{a}_2^T$. Vi bestämmer en bas i detta nollrum på vanligt sätt,

$$\begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix},$$

och ser att $\{(-1 \quad 1 \quad 1)^T\}$ är en sådan.

Exempel 5.16. Om underrummet W är definierat som mängden av alla $\mathbf{x} = (x_1 \quad x_2 \quad x_3 \quad x_4)^T \in \mathbf{R}^4$ som satisfierar systemet

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

så består W av alla $\mathbf{x}$, som är ortogonala mot $(1 \quad 1 \quad 1 \quad 1)^T$ och $(1 \quad 2 \quad 3 \quad 4)^T$, dvs. dessa två vektorer uppspänner $W^\perp$. Eftersom de råkar vara linjärt oberoende, bildar de en bas i $W^\perp$.

Sats 5.5 har en föjdsats som direkt berör ekvationslösning:

Sats 5.6. *För en ekvation $A\mathbf{x} = \mathbf{b}$ gäller att för varje $\mathbf{b} \in \mathbf{R}(A)$ finns i mängden av alla lösningar* precis en *lösning $\mathbf{x}$ i radrummet $R(A^T)$.*

Bevis. Eftersom $\mathbf{b} \in R(A)$, existerar det lösningar. Låt $\mathbf{y}$ vara någon lösning och skriv (med stöd av Sats 5.5 och Sats 5.4) denna som en summa

$$\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2\,, \text{ där } \mathbf{y}_1 \in R(A^T) \text{ och } \mathbf{y}_2 \in N(A)\,.$$

Då är $A\mathbf{y}_2 = \mathbf{0}$ så att

$$\mathbf{b} = A\mathbf{y} = A\mathbf{y}_1 + A\mathbf{y}_2 = A\mathbf{y}_1\,,$$

vilket bevisar existensen av en lösning $\mathbf{x} = \mathbf{y}_1 \in R(A^T)$. Det återstår att bevisa entydigheten hos detta $\mathbf{x}$. Låt $\mathbf{x}'$ beteckna en godtycklig lösning i $R(A^T)$. Då är

$$A(\mathbf{x} - \mathbf{x}') = A\mathbf{x} - A\mathbf{x}' = \mathbf{b} - \mathbf{b} = \mathbf{0}\,,$$

dvs. $\mathbf{x} - \mathbf{x}' \in N(A)$. Men $\mathbf{x} - \mathbf{x}'$ ligger ju också i det ortogonala komplementet $R(A^T)$ till $N(A)$. Således är vektorn $\mathbf{x} - \mathbf{x}'$ ortogonal mot sig själv. Men den enda vektor som kan vara ortogonal mot sig själv är nollvektorn (axiom IV (a)). Alltså är $\mathbf{x}' = \mathbf{x}$. $\diamondsuit$

Exempel 5.17. För att bestämma den lösning till

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

som ligger i koefficientmatrisens radrum $R(A^T)$, löser vi först systemet och finner lösningarna

$$\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} = \mathbf{x}_0 + s\mathbf{v}_1 + t\mathbf{v}_2\,.$$

Eftersom $R(A^T)$ är ortogonalt mot $N(A) = \text{spn}\{\mathbf{v}_1, \mathbf{v}_2\}$, bör vi välja s och t så att

$$\begin{cases} 0 = \mathbf{v}_1^T\mathbf{x} = \mathbf{v}_1^T\mathbf{x}_0 + s\|\mathbf{v}_1\|^2 + t\mathbf{v}_1^T\mathbf{v}_2 \\ 0 = \mathbf{v}_1^T\mathbf{x} = \mathbf{v}_2^T\mathbf{x}_0 + s\mathbf{v}_2^T\mathbf{v}_1 + t\|\mathbf{v}_2\|^2 . \end{cases}$$

Efter uträkning av skalära produkter och normer fås systemet

$$\begin{cases} 0 = -2 + 6s + 8t \\ 0 = -3 + 8s + 14t , \end{cases}$$

som har lösningen $s = 1/5$, $t = 1/10$. Insättning av dessa värden ger den lösningsvektor $\mathbf{x} = \frac{1}{10}\begin{pmatrix} 4 & 3 & 2 & 1 \end{pmatrix}^T$, som ligger i $R(A^T)$.

Kolonnrum och nollrum för en matrisprodukt

Sats 5.7. *Antag att A är en m/n-matris och B en n/p-matris. Då gäller:*

$$\begin{array}{ll} (i) & R(AB) \subseteq R(A) , \\ (ii) & N(AB) \supseteq N(B) . \end{array}$$

Bevis. (i) Formeln följer ur implikationerna

$$\begin{aligned} \mathbf{b} \in R(AB) \quad &\Longrightarrow \quad \mathbf{b} = AB\mathbf{x} \text{ för något } \mathbf{x} \in \mathbf{R}^p \\ &\Longrightarrow \quad \mathbf{b} = A\mathbf{y} \text{ för något } \mathbf{y} \in \mathbf{R}^n \text{ (nämligen } \mathbf{y} = B\mathbf{x}) \\ &\Longrightarrow \quad \mathbf{b} \in R(A) . \end{aligned}$$

(ii) Formeln fås ur implikationerna

$$\mathbf{x} \in N(B) \quad \Longrightarrow \quad B\mathbf{x} = \mathbf{0} \quad \Longrightarrow \quad AB\mathbf{x} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{x} \in N(AB) . \diamondsuit$$

Anmärkning. Om vi tillämpar formlerna i Sats 5.7 på produkten $B^TA^T = (AB)^T$, fås formler som gäller radrummet och vänsternollrummet av en produkt:

$$\begin{aligned} R((AB)^T) &\subseteq R(B^T) , \\ N((AB)^T) &\supseteq N(A^T) . \end{aligned}$$

Låt $r(X)$ och $d(X)$ beteckna rangen $\dim R(X)$ respektive defekten $\dim N(X)$ hos en matris X. Låt dessutom $\min(\alpha, \beta)$ beteckna det mindre av talen α och β (utläses: "minimum av α och β").

Korollarium 5.7. *Om A är en m/n-matris och B en n/p-matris så är*

$$\begin{array}{ll} (i) & r(AB) \leq \min(r(A), r(B)) , \\ (ii) & d(AB) \geq d(B) . \end{array}$$

Bevis. Relation (ii) är en direkt följd av Sats 5.7. Samma sats säger också direkt att $r(AB) \leq r(A)$. Med hjälp av Sats 4.13 och anmärkningen ovan, fås

$$r(AB) = \dim R((AB)^T) \leq \dim R(B^T) = \dim R(B) = r(B) .$$

Alltså är talet $r(AB)$ mindre än både $r(A)$ och $r(B)$, dvs. (i) gäller. $\diamondsuit$

Submatriser

Definition 5.8. Om A är en given matris, så fås en *submatris* av A genom att ett antal rader och kolonner stryks i A.

Exempel 5.18. Om man i matrisen

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

stryker rad ett och tre samt kolonn två, så uppstår submatrisen $\begin{pmatrix} 4 & 6 \end{pmatrix}$.

Sats 5.8. *Antag att $r = r(A)$ är rangen hos en matris A. Då gäller:*

$$\begin{array}{ll} (i) & r(C) \leq r \text{ för varje submatris } C \text{ av } A; \\ (ii) & r(C) = r \text{ för någon } r/r\text{-submatris } C \text{ av } A. \end{array}$$

Bevis. (i) Vi "reducerar" A till C i två steg. Först stryker vi rader i A så att en matris B uppstår. Då kan ju dimensionen för radrummet bara minska. Enligt Sats 4.13 är därför

$$r = \dim R(A) = \dim R(A^T) \geq \dim R(B^T) = \dim R(B) = r(B)\,.$$

I det andra steget stryker vi kolonner i B så att C uppstår. Då kan ju dimensionen för kolonnrummet bara minska, så att

$$r(B) = \dim R(B) \geq \dim R(C) = r(C)\,.$$

Alltså är $r(C) \leq r$.

(ii) Precis r rader i A, säg $\bar{\mathbf{a}}_1, \ldots, \bar{\mathbf{a}}_r$, bildar en bas i $R(A^T)$. Sätt

$$B = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_r \end{pmatrix}.$$

Då är $r(B) = \dim R(B^T) = r$. Följaktligen bildar precis r kolonner i B, säg $\mathbf{b}_1, \ldots, \mathbf{b}_r$, en bas i $R(B)$. Sätt $C = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_r \end{pmatrix}$. Då är $r(C) = r$ och C är en r/r-submatris av matrisen A. $\diamondsuit$

Övningsuppgifter

1. Vektorerna $\mathbf{a}$ och $\mathbf{b}$ har längderna 2 respektive 3 och vinkeln mellen dem är $2\pi/3$. (a) Beräkna längden av vektorn $\mathbf{v} = 2\mathbf{a} + 3\mathbf{b}$. (b) Bestäm någon vektor av formen $x\mathbf{a} + y\mathbf{b}$, som är ortogonal mot $\mathbf{v}$.
2. Bestäm a så att vinkeln mellan vektorerna $\begin{pmatrix} 3 & 4 & 5a \end{pmatrix}^T$ och $\begin{pmatrix} 3a & 4a & 5 \end{pmatrix}^T$ blir $\pi/3$.
3. Beräkna vinkeln mellan vektorerna $\mathbf{a}$ och $\mathbf{b}$ ($\mathbf{a} \neq \mathbf{0}$, $\mathbf{b} \neq \mathbf{0}$) om $\mathbf{a} + 3\mathbf{b}$ är vinkelrät mot $2\mathbf{a} - \mathbf{b}$ och $\mathbf{a} + 7\mathbf{b}$ är vinkelrät mot $2\mathbf{a} + \mathbf{b}$.

4. Bestäm en enhetsvektor som är parallell med xy-planet och vinkelrät mot vektorn $(\,4 \quad -3 \quad 1\,)^T$.
5. Bevisa att $x^2+y^2+z^2 \geq 6$ om $2x-y+z \leq -6$. Ledning: Använd Schwarz olikhet.
6. Bevisa att $(a_1 + a_2 + \cdots + a_n)^2 \leq n({a_1}^2 + {a_2}^2 + \cdots + {a_n}^2)$. Ledning: Använd Schwarz olikhet.
7. Visa att om $\mathbf{x} \in \mathbf{R}^n$ och $\mathbf{x}^T\mathbf{y} = 0$ för alla $\mathbf{y} \in \mathbf{R}^n$ så måste $\mathbf{x}$ vara nollvektorn.
8. Visa att om $N(AB) = N(B)$ så är $R(B) \cap N(A) = \{\mathbf{0}\}$.
9. Betrakta följande underrum av $\mathbf{R}^4$:

$$\begin{aligned}
M_1 &= \operatorname{spn}\{(\,1 \quad -1 \quad 1 \quad -1\,)\}\,,\\
M_2 &= \operatorname{spn}\{(\,1 \quad 0 \quad 0 \quad 1\,), (\,0 \quad 1 \quad 0 \quad -1\,)\}\,,\\
M_3 &= \operatorname{spn}\{(\,1 \quad 0 \quad 0 \quad -1\,), (\,0 \quad 1 \quad 0 \quad -1\,), (\,0 \quad 0 \quad 1 \quad 1\,)\}\,,\\
M_4 &= \operatorname{spn}\{(\,-1 \quad 1 \quad 0 \quad 1\,), (\,0 \quad 0 \quad 1 \quad 0\,)\}\,.
\end{aligned}$$

   (a) Ange alla par av ortogonala underrum av $\mathbf{R}^4$ i mängden $\{M_1, M_2, M_3, M_4\}$.

   (b) Undersök vilka av dessa par som är varandras ortogonala komplement.
10. Visa att om A och B är n/n-matriser och $\mathbf{x}$ och $\mathbf{y}$ är två vektorer i $\mathbf{R}^n$ så är skalära produkten av $A\mathbf{x}$ och $B\mathbf{y}$ likamed skalära produkten av $\mathbf{x}$ och $(A^TB)\mathbf{y}$.
11. Bestäm i $\mathbf{R}^3$ alla vektorer som är ortogonala mot både $(\,1 \quad 1 \quad 1\,)$ och $(\,1 \quad -1 \quad 0\,)$. Konstruera utgående från detta ett system av tre parvis ortogonala enhetsvektorer i $\mathbf{R}^3$ (dvs. ett ortonormerat system).
12. Rad i i matrisen B, som antas inverterbar, är ortogonal mot kolonn j i B^{-1} om $i \neq j$. Varför?
13. Visa att om V och W är ortogonala underrum av ett euklidiskt vektorrum så är nollvektorn den enda vektor de har gemensam.
14. Finns det något värde på k, för vilket det existerar en rät linje $x = at$, $y = bt$, $z = ct$ som är parallell med de tre planen

$$\begin{aligned}
2x - (k+2)y + z + 1 &= 0\,,\\
kx + 2y - 4z - 3 &= 0\,,\\
kx - y - 3z + 3 &= 0\,.
\end{aligned}$$

   Ange i så fall k och vektorn $(\,a \quad b \quad c\,)^T$.
15. Konstruera en matris, vars nollrum spänns upp av vektorn $(\,1 \quad 0 \quad 1\,)^T$.
16. (a) Låt W vara det underrum av $\mathbf{R}^5$ som uppspänns av $\mathbf{u} = (\,1 \quad 2 \quad 3 \quad -1 \quad 2\,)^T$ och $\mathbf{v} = (\,2 \quad 4 \quad 7 \quad 2 \quad -1\,)^T$. Bestäm en bas i det ortogonala komplementet $W^\perp$ till W.

   (b) Låt U vara det underrum av $\mathbf{R}^3$ som uppspänns av $(\,1 \quad 1 \quad -2\,)^T$. Bestäm en bas i $U^\perp$ samt skriv vektorn $\mathbf{x} = (\,2 \quad 1 \quad -3\,)^T$ som en summa $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, där $\mathbf{x}_1 \in U$ och $\mathbf{x}_2 \in U^\perp$.

17. (*a*) Visa att vektorn $\mathbf{b} = (\,9 \quad -6 \quad 3\,)^T$ tillhör kolonnrummet för matrisen

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

(*b*) Beräkna den entydiga vektor $\mathbf{x}$ i radrummet för A, för vilken $A\mathbf{x} = \mathbf{b}$.

18. Visa att om $\{\mathbf{x},\,\mathbf{y}\}$ är ett ON-system i ett euklidiskt vektorrum så är $\|\mathbf{x} - \mathbf{y}\| = \sqrt{2}$.

19. Antag att V spänns upp av vektorerna $(\,1 \quad 1 \quad 0 \quad 1\,)^T$ och $(\,1 \quad 2 \quad 0 \quad 0\,)^T$. Bestäm ett underrum W av $\mathbf{R}^4$ sådant att $V \cap W = \{\mathbf{0}\}$ och $V + W = \mathbf{R}^4$ (se uppgift 36 i kapitel 4).

20. (*a*) Visa att i ett euklidiskt vektorrum gäller att

$$(\mathbf{x},\,\mathbf{y}) = \frac{1}{4}\left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\right)$$

för varje $\mathbf{x}$ och $\mathbf{y}$.

(*b*) Visa att i ett euklidiskt vektorrum gäller att

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2,$$

för varje $\mathbf{x}$ och $\mathbf{y}$. Tolka denna formel geometriskt!

21. Visa med hjälp av uppgift 26 i kapitel 3 att

$$(\mathbf{x},\,\mathbf{y}) = \mathbf{x}^T \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix} \mathbf{y}$$

uppfyller axiom IV (*a*) för en skalär produkt (de övriga axiomen följer ur allmänna räkneregler rörande matrisprodukter).

22. Låt $X = \{x_1, \ldots, x_n\}$ vara en mängd ($n \geq 3$) och låt $A_1, \ldots, A_m$ vara äkta delmänger av X (då är $A_k \neq X$ för alla k), sådana att varje par x_i, x_j av olika element i X bägge finns i precis en delmängd A_k. Låt $B = (b_{ik})$ vara den s.k. *incidensmatrisen*, som är definierad av att $b_{ik} = 1$ om $x_i \in A_k$ och $b_{ik} = 0$ om $x_i \notin A_k$

(a) Visa att BB^T är en matris som består av enbart ettor utom i diagonalen, där vi har talen $r_1, \ldots, r_n$. Dessa tal r_i anger för hur många k det gäller att $x_i \in A_k$ (inte sant?).

(b) Notera att $r_i \geq 2$ på grund av att $n \geq 3$.

(c) Matrisen BB^T är icke-singulär enligt uppgift 26 i kapitel 3. Använd formeln för rangen av en matrisprodukt till att visa att $m \geq n$, dvs. använd formeln

$$r(CD) \leq \min(r(C), r(D)).$$

6. Determinanter

Innan vi slår fast en definition av begreppet determinant, behöver vi vissa förberedande förklaringar:

En *permutation* av talen $1, \ldots, n$ är en uppräkning $(j_1, \ldots, j_n)$ av dessa samma tal i någon ordning. Mängden av alla permutationer av $1, \ldots, n$ betecknar vi med S_n. Observera att en permutation $\sigma = (j_1, \ldots, j_n)$ också kan uppfattas som en funktion $\sigma : k \mapsto \sigma(k) = j_k$ med indexet k som argument och talet j_k som funktionsvärde.

Exempel 6.1. Både $(2, 1, 4, 3)$ och $(4, 3, 1, 2)$ är permutationer av talen 1, 2, 3, 4.

En *inversion* förekommer vid (i, k) (och (k, i)) i en permutation $\sigma = (j_1, \ldots, j_n)$ om för två index i och k med $i < k$ gäller $j_i > j_k$, dvs. om talen j_i och j_k kommer i omvänd storleksordning. Det totala antalet inversioner i $\sigma = (j_1, \ldots, j_n)$ är

$$\begin{aligned} t(\sigma) = \quad & \#(k > 1 \quad \text{med} \quad j_1 > j_k) \\ & + \#(k > 2 \quad \text{med} \quad j_2 > j_k) + \cdots + \#(k > n-1 \quad \text{med} \quad j_{n-1} > j_k)\,. \end{aligned}$$

Permutationen σ är *jämn* om $t(\sigma)$ är ett jämnt tal och *udda* om $t(\sigma)$ är udda.

Exempel 6.2. Om $\sigma = (4, 3, 1, 2)$ så är $t(\sigma) = 3 + 2 + 0 = 5$, dvs. σ är en udda permutation.

Låt nu $A = (a_{ij})$ vara en n/n-matris.

Definition 6.1. En *elementär produkt* i A är en produkt av n element i A, vilka alla tagits ur **olika** rader och kolonner.

Exempel 6.3. Om

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

så är t.ex. $a_{12}a_{21}a_{33}$ och $a_{11}a_{23}a_{32}$ elementära produkter i A.

Definition 6.2. *Determinanten* av matrisen $A = (a_{ik})$ är summan

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{1j_1} a_{2j_2} \cdots a_{nj_n}\,,$$

(där $\sigma = (j_1, \ldots, j_n)$) av alla de $n!$ elementära produkterna i A försedda med tecknet

$$\epsilon(\sigma) = (-1)^{t(\sigma)} = \begin{cases} 1, & \text{om } \sigma \text{ är en jämn permutation} \\ -1, & \text{om } \sigma \text{ är en udda permutation.} \end{cases}$$

Oftast använder man beteckningen

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ - & - & - & - \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

för determinanten $\det(A)$.

Exempel 6.4. Matrisen

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

har de två elementära produkterna $a_{11}a_{22}$ och $a_{12}a_{21}$. Den första av dessa får ett positivt tecken, eftersom permutationen $(1, 2)$ är jämn, medan den andra får ett negativt tecken, eftersom permutationen $(2, 1)$ är udda. Alltså är

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

Exempel 6.5. För en 3/3-matris

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

har vi $3! = 6$ elementära produkter att bestämma tecknet för:

Elementär produkt	$\epsilon(\sigma)$
$a_{11}a_{22}a_{33}$	$+1$
$a_{11}a_{23}a_{32}$	-1
$a_{12}a_{21}a_{33}$	-1
$a_{12}a_{23}a_{31}$	$+1$
$a_{13}a_{21}a_{32}$	$+1$
$a_{13}a_{22}a_{31}$	-1

T.ex. får den elementära produkten i rad två ett minustecken, eftersom permutationen $(1, 3, 2)$ är udda. Tecknen i denna tabell kan sammanfattas i *Sarrus regel*: Upprepa A:s två första kolonner efter A:

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

De tre produkter som bildas av element på en (åt höger) snett nedåtgående diagonallinje genom denna rektangel av tal, är de elementära produkter som får plustecken (t.ex. huvuddiagonalen i A) medan de tre produkter, som bildas av element på en snett uppåtgående diagonallinje, är de elementära produkter som får minustecken. Kom ihåg att **Sarrus regel bara gäller för 3/3-matriser**.

Egenskaper hos determinanter

En n/n-determinant ger upphov till $n!$ elementära produkter. Om n är stort (ja, redan om $n = 4$) innehåller definitionsuttrycket för en determinant alltför många termer för att det skall löna sig att använda detta då man räknar ut determinantens värde. De egenskaper hos determinanter som vi nu skall härleda utgör i själva verket **räkneregler** som kommer att ge oss effektiva metoder att räkna ut determinanter – både stora och små.

1. *En n/n-matris och dess transponerade matris har samma determinant*

$$\det(A^T) = \det(A)\,.$$

Bevis. Då man bildar en elementär produkt, tar man exakt ett element ur varje rad och kolonn. Alltså ger A och A^T upphov till precis samma elementära produkter. Det återstår för oss bara att förklara varför dessa elementära produkter förses med samma tecken. I en godtycklig elementär produkt $a_{1j_1} \cdots a_{nj_n}$ i A kan vi byta ordningsföljd på faktorerna så att den kan tolkas som en elementär produkt för A^T:

$$a_{1j_1} \cdots a_{nj_n} = a_{k_1 1} \cdots a_{k_n n}\,.$$

Då är permutationerna $\sigma : i \mapsto j_i$ och $\tau : r \mapsto k_r$ varandras inversa funktioner, $\tau = \sigma^{-1}$. En inversion förekommer i σ vid (i, l) om och endast om $l - i$ och $\sigma(l) - \sigma(i)$ har olika tecken, dvs. om och endast om

$$\frac{\sigma(l) - \sigma(i)}{l - i} = \frac{j_l - j_i}{l - i} < 0\,.$$

Denna kvot är negativ om och endast om den inverterade kvoten är negativ,

$$\frac{\tau(j_l) - \tau(j_i)}{j_l - j_i} = \frac{l - i}{j_l - j_i} < 0\,,$$

och detta gäller precis då det förekommer en inversion i τ vid (j_l, j_i). Permutationerna σ och τ innehåller alltså lika många inversioner och ger upphov till samma tecken. $\diamondsuit$

Anmärkning. På grund av egenskap 1 har rader och kolonner i en determinant samma ställning. Därför **kan räknereglerna 2 till 6**, som vi nedan formulerar för kolonner, utan vidare **formuleras också för rader**.

2. *Låt $A = (\,\mathbf{a}_1 \;\; \ldots \;\; \mathbf{a}_i \;\; \ldots \;\; \mathbf{a}_k \;\; \ldots \;\; \mathbf{a}_n\,)$ vara en matris och låt A' vara den matris som uppkommer då kolonnerna $\mathbf{a}_i$ och $\mathbf{a}_k$ byter plats i A. Då är*

$$\det(A') = -\det(A)\,.$$

En motsvarande regel gäller vid ombyte av rader.

Bevis. Låt $\sigma = (j_1, \ldots, j_i, \ldots, j_k, \ldots, j_n)$ beteckna en en godtycklig permutation och låt σ' vara den permutation som uppstår då j_i och j_k byter plats i σ. Vi kan visa att $t(\sigma) - t(\sigma') = u$ alltid är ett udda tal. Därför är

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1j_1} \cdots a_{ij_i} \cdots a_{kj_k} \cdots a_{nj_n} \\
&= \sum_{\sigma' \in S_n} -\epsilon(\sigma') a_{1j_1} \cdots a_{kj_k} \cdots a_{ij_i} \cdots a_{nj_n} = -\det(A')\,.
\end{aligned}$$

Resonemanget är litet komplicerat men vi återger det för fullständighetens skull.

Bevis för att u är udda: Bara inversioner av typerna

$$\begin{array}{lll} (1) & j_i > j_l, & i < l < k\,, \\ (2) & j_l > j_k, & i < l < k\,, \end{array}$$

kan uppkomma eller försvinna då j_i och j_k byter plats i σ. Låt I vara mängden av alla heltal l mellan i och k och sätt

$$\begin{aligned} n_1 &= \#(l \in I \text{ med } j_i < j_l), & m_1 &= \#(l \in I \text{ med } j_l < j_k), \\ n_2 &= \#(l \in I \text{ med } j_i > j_l), & m_2 &= \#(l \in I \text{ med } j_l > j_k)\,. \end{aligned}$$

Antalet inversioner förändras med

$u_1 = n_1 - n_2$, pga. att j_i flyttas till andra sidan av j_l,
$u_2 = m_1 - m_2$, pga. att j_k flyttas till andra sidan av j_l,
$u_3 = \pm 1$, pga. att j_i och j_k byter plats.

Eftersom $n_1 + n_2 = m_1 + m_2 = (k - i) - 1$ är konstant, så förändras u_1 och u_2 med steglängden ± 2 då n_1, n_2 respektive m_1, m_2 ändras. Således är antingen u_1 och u_2 bägge udda (då $(k-i)-1$ är udda) eller bägge jämna (då $(k-i)-1$ är jämnt). Men då är ju $u_1 + u_2$ alltid jämnt och den totala förändringen av antalet inversioner $u = u_1 + u_2 + u_3 = u_1 + u_2 \pm 1$ alltid udda. $\diamondsuit$

3. *Om två kolonner* (eller rader) *är lika i matrisen A, så är*

$$\det(A) = 0\,.$$

Bevis. Om $A = (\ldots\ \mathbf{a}_i\ \ldots\ \mathbf{a}_k\ \ldots)$ och $\mathbf{a}_i = \mathbf{a}_k$, så är enligt regel 2

$$\det(A) = \det(\ldots\ \mathbf{a}_i\ \ldots\ \mathbf{a}_k\ \ldots) = -\det(\ldots\ \mathbf{a}_k\ \ldots\ \mathbf{a}_i\ \ldots) = -\det(A)$$

och därmed $\det(A) = 0$. $\diamondsuit$

Nästa egenskap följer direkt ur definitionen på en determinant och kräver därför inget bevis:

4. *Om en kolonn* (eller rad) *innehåller en gemensam faktor, så kan denna faktor flyttas ut ur determinanten:*

$$\det(\mathbf{a}_1\ \ldots\ \lambda\mathbf{a}_i\ \ldots\ \mathbf{a}_n) = \lambda \det(\mathbf{a}_1\ \ldots\ \mathbf{a}_i\ \ldots\ \mathbf{a}_n)\,.$$

Exempel 6.6. I nedanstående determinant tar vi ut faktorn 2 ur första raden och faktorn 3 ur andra raden och får värdet

$$\begin{vmatrix} 2 & 2 & 4 \\ 3 & 3 & 6 \\ 1 & 9 & 8 \end{vmatrix} = 2 \cdot 3 \cdot \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 9 & 8 \end{vmatrix} = 2 \cdot 3 \cdot 0 = 0\,,$$

eftersom de två första raderna blir lika (regel 3).

5. *En determinant är additiv i varje kolonn* (och rad),

$$\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_i + \mathbf{a}'_i \ \dots \ \mathbf{a}_n) = \det(\mathbf{a}_1 \ \dots \ \mathbf{a}_i \ \dots \ \mathbf{a}_n) + \det(\mathbf{a}_1 \ \dots \ \mathbf{a}'_i \ \dots \ \mathbf{a}_n)\,.$$

Bevis. Enligt distributionslagen för tal kan de elementära produkterna i vänstra ledet skrivas

$$a_{1j_1} \cdots (a_{ij_i} + a'_{ij_i}) \cdots a_{nj_n} = a_{1j_1} \cdots a'_{ij_i} \cdots a_{nj_n} + a_{1j_1} \cdots a_{ij_i} \cdots a_{nj_n}\,.$$

Regeln följer ur detta och definitionen på en determinant. $\diamondsuit$

Exempel 6.7. Om vi tillämpar regel 5 på rader får vi att

$$\begin{vmatrix} 7 & 9 & 5 \\ 2 & 1 & 3 \\ 2 & 2 & 1 \end{vmatrix} + \begin{vmatrix} 7 & 9 & 5 \\ 3 & 0 & 1 \\ 2 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 7 & 9 & 5 \\ 2+3 & 1+0 & 3+1 \\ 2 & 2 & 1 \end{vmatrix}\,.$$

Som en direkt följd av reglerna 5, 4 och 3 fås:

6. *Om $i \neq k$ så är*

$$\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_i \ \dots \ \mathbf{a}_n) = \det(\mathbf{a}_1 \ \dots \ \mathbf{a}_i - \lambda\mathbf{a}_k \ \dots \ \mathbf{a}_n)\,, \qquad (i \neq k)\,.$$

Anmärkning. Formulerad för rader säger denna regel att en determinants värde förblir oförändrat vid användning av $(BO1)$:

$$\det\begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_i \\ \vdots \\ \bar{\mathbf{a}}_n \end{pmatrix} = \det\begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_i - \lambda\bar{\mathbf{a}}_k \\ \vdots \\ \bar{\mathbf{a}}_n \end{pmatrix}\,, \qquad (i \neq k)\,.$$

Om vi således med en följd av användningar av $(BO1)$ kan överföra en matris A på en uppåt triangulär form U,

$$A \to \cdots \to U = \begin{pmatrix} u_{11} & * & \cdots & * \\ 0 & u_{22} & \cdots & * \\ - & - & - & - \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}\,,$$

så är $\det(A) = \det(U) = u_{11}u_{22}\cdots u_{nn}$. I en triangulär matris är nämligen produkten av diagonalelementen den enda elementära produkt som kan vara olik noll. Alla andra elementära produkter innehåller minst en nolla.

Detta ger oss en effektiv metod att räkna ut värdet på en determinant. Reglerna 2 och 4 utgör ju något modifierade varianter av $(BO2)$ och $(BO3)$, så i praktiken **har vi Gausselimineringens alla verktyg till vårt förfogande**:

Exempel 6.8. För att kunna använda det "enkla" pivotelementet 1 flyttar vi i nedanstående determinant först ut den gemensamma faktorn 2 i andra raden och låter samtidigt de två första raderna byta plats … :

$$\begin{vmatrix} 3 & 1 & 1 \\ 2 & 4 & 6 \\ 2 & 1 & 0 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 1 \\ 2 & 1 & 0 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & 3 \\ 0 & -5 & -8 \\ 0 & -3 & -6 \end{vmatrix}$$
$$= -(-2)(-3) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -5 & -8 \end{vmatrix} = -6 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{vmatrix} = -6 \cdot 1 \cdot 1 \cdot 2 = -12\,.$$

7. *En kvadratisk matris A är singulär om och endast om* $\det(A) = 0$.

Bevis. Matrisen A kan överföras på en uppåt triangulär matris U med hjälp av enbart $(BO1)$ och $(BO2)$. Nu gäller

$$\begin{aligned} A \text{ är singulär} &\iff U \text{ är singulär} \\ &\iff \text{något diagonalelement i } U \text{ är noll} \\ &\iff \det(U) = 0 \\ &\iff \det(A) = 0\,, \end{aligned}$$

ty $\det(A) = \pm\det(U)$, där vi har ett minustecken om $(BO2)$ har använts ett udda antal gånger.

8. *Om A och B är n/n-matriser, så är*

$$\det(AB) = \det(A) \cdot \det(B)\,.$$

Determinanter kan alltså multipliceras med varandra på samma sätt som man multiplicerar matriser.

Bevis. (a) Vi visar först att formeln gäller om A är en diagonalmatris. Om

$$A = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \quad \text{och} \quad B = \begin{pmatrix} \overline{\mathbf{b}}_1 \\ \vdots \\ \overline{\mathbf{b}}_n \end{pmatrix},$$

så är enligt regel 4

$$\det(AB) = \det \begin{pmatrix} d_1\overline{\mathbf{b}}_1 \\ \vdots \\ d_n\overline{\mathbf{b}}_n \end{pmatrix} = d_1 \cdots d_n \cdot \det \begin{pmatrix} \overline{\mathbf{b}}_1 \\ \vdots \\ \overline{\mathbf{b}}_n \end{pmatrix} = \det(A) \cdot \det(B)\,.$$

(b) Antag att A är en godtycklig n/n-matris. Om A **är singulär**, så är också A^T singulär och följaktligen också $B^T A^T = (AB)^T$. Nu är både $\det(A) = 0$ och $\det(AB) = \det((AB)^T) = 0$ enligt regel 7. Formeln gäller alltså i detta fall, eftersom bägge leden är noll.

Om A **inte är singulär**, så kan A överföras på en diagonalmatris D med hjälp av enbart $(BO1)$ och $(BO2)$:

$$A \to \cdots \to D = \text{ diagonalmatris}, \tag{1}$$

där $\det(A) = \pm\det(D)$. Det första steget i (1) får i det fortsatta resonemanget representera vilket steg som helst. Om detta är $(BO1)$-operationen

$$A = \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_i \\ \vdots \\ \bar{\mathbf{a}}_n \end{pmatrix} \to \begin{pmatrix} \bar{\mathbf{a}}_1 \\ \vdots \\ \bar{\mathbf{a}}_i - \lambda\bar{\mathbf{a}}_k \\ \vdots \\ \bar{\mathbf{a}}_n \end{pmatrix} = A_1\,,$$

så fås med samma basoperation och samma λ, att

$$AB = \begin{pmatrix} \bar{\mathbf{a}}_1 B \\ \vdots \\ \bar{\mathbf{a}}_i B \\ \vdots \\ \bar{\mathbf{a}}_n B \end{pmatrix} \to \begin{pmatrix} \bar{\mathbf{a}}_1 B \\ \vdots \\ \bar{\mathbf{a}}_i B - \lambda\bar{\mathbf{a}}_k B \\ \vdots \\ \bar{\mathbf{a}}_n B \end{pmatrix} = A_1 B\,.$$

Är det första steget i (1) igen en $(BO2)$-operation

$$A = \begin{pmatrix} \vdots \\ \bar{\mathbf{a}}_i \\ \vdots \\ \bar{\mathbf{a}}_k \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ \bar{\mathbf{a}}_k \\ \vdots \\ \bar{\mathbf{a}}_i \\ \vdots \end{pmatrix} = A_1\,,$$

så fås med ombyte av samma rader

$$AB = \begin{pmatrix} \vdots \\ \bar{\mathbf{a}}_i B \\ \vdots \\ \bar{\mathbf{a}}_k B \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ \bar{\mathbf{a}}_k B \\ \vdots \\ \bar{\mathbf{a}}_i B \\ \vdots \end{pmatrix} = A_1 B\,.$$

Precis samma sekvens av basoperationer som i (1), använd på produkten AB, leder alltså till

$$AB \to \cdots \to DB\,,$$

där det enligt tidigare regler och enligt fall (a) gäller att

$$\det(AB) = \pm\det(DB) = \pm\det(D)\cdot\det(B) = \det(A)\cdot\det(B)\,.$$

Formeln i regel 8 gäller alltså generellt för godtyckliga matriser A och B. $\diamondsuit$

Utveckling efter en rad eller en kolonn

Termerna i definitionsuttrycket för en determinant kan grupperas så att varje grupp representer ett element i determinanten multiplicerat med determinanten för en submatris försedd med plus- eller minustecken. Vi illustrerar detta med hjälp av en 3/3-determinant:

Exempel 6.9. Med hjälp av definitionen (eller Sarrus regel) utvecklar vi determinanten samt grupperar ihop de termer som vi nedan har skrivit under varandra:

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

Det nästsista och sista uttrycket kallas utvecklingen av determinanten efter första raden. Genom att gruppera termerna på annat sätt får man utvecklingar efter andra rader eller kolonner.

Den “subdeterminant med tecken” som vi i ovanstående exempel betecknat med A_{ik} kallas *kofaktorn* (eller *komplementet*) till elementet a_{ik}. Allmänt får vi:

Sats 6.1. *Om $A = (a_{ik})$ är en n/n-matris och $\hat{A}_{ik}$ betecknar den submatris som uppstår då rad i och kolonn k stryks i A, så är*

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

utvecklingen av $\det(A)$ efter rad i, *där* kofaktorerna A_{ik} *fås ur*

$$A_{ik} = (-1)^{i+k}\det(\hat{A}_{ik})\,.$$

På motsvarande sätt kan $\det(A)$ *utvecklas efter en kolonn.*

Tecknet på en kofaktor bestäms enklast genom att man kommer ihåg att matriselementet längst uppe till vänster är associerat med en kofaktor med ett plustecken framför en subdeterminant. De övriga elementen är sedan omväxlande associerade med plus- och minustecken precis som rutorna på ett schackbräde omväxlande är vita och svarta.

Exempel 6.10. Om vi vill utveckla följande 3/3-determinant efter 2:a kolonnen, så bör vi alltså associera ett minustecken med elementen 6 och 2 och ett plustecken med elementet -1 i andra kolonnen. Vi får utvecklingen:

$$\begin{vmatrix} 4 & 6 & 3 \\ 3 & -1 & 7 \\ 7 & 2 & 5 \end{vmatrix} = -6\begin{vmatrix} 3 & 7 \\ 7 & 5 \end{vmatrix} - \begin{vmatrix} 4 & 3 \\ 7 & 5 \end{vmatrix} - 2\begin{vmatrix} 4 & 3 \\ 3 & 7 \end{vmatrix}\,.$$

Matrisinverser

Betrakta utvecklingen efter rad k av determinanten för en n/n-matris $A = (a_{lm})$:

$$\det(A) = a_{k1}A_{k1} + \cdots + a_{kn}A_{kn}\,. \tag{2}$$

Den likartade summan

$$a_{i1}A_{k1} + \cdots + a_{in}A_{kn}\,, \qquad (i \neq k)\,,$$

kan uppenbarligen betraktas som utvecklingen efter rad k av den determinant som vi får ur $\det(A)$ genom att ersätta rad k med en kopia av rad i. Eftersom två rader är lika i denna determinant, är dess värde noll:

$$a_{i1}A_{k1} + \cdots + a_{in}A_{kn} = 0\,, \qquad (i \neq k)\,. \tag{3}$$

Likheterna (2) och (3) kan skrivas på en enda rad med hjälp av Kroneckers delta,

$$a_{i1}A_{k1} + \cdots + a_{in}A_{kn} = \det(A) \cdot \delta_{ik}\,.$$

Vänstra ledet kan nu tolkas som ett matriselement i produkten av A med

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ - & - & - & - \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

vilken ibland kallas den *adjungerade matrisen* till A. Observera att det första indexet här är kolonnindex medan det andra är radindex. Vi har alltså likheten $A\tilde{A} = \det(A)\cdot I$, vilket ger formeln

$$A^{-1} = \frac{1}{\det(A)}\tilde{A}$$

för matrisinversen om $\det(A) \neq 0$ (dvs. om A är icke-singulär).

Anmärkning. Denna formel är intressant ur teoretisk synvinkel men olämplig för numerisk uträkning av matrisinverser. Alla de n^2 matriselementen är ju $(n-1)/(n-1)$-determinanter, som var och en skall räknas ut separat.

Exempel 6.11. Om $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, så är

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\,,$$

ifall $\det(A) = ad - bc \neq 0$.

Cramers regel

Om n/n-matrisen A är inverterbar så kan lösningen till en ekvation $A\mathbf{x} = \mathbf{b}$ skrivas

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)}\tilde{A}\mathbf{b}\,,$$

vilket i komponentform ger

$$x_j = \frac{1}{\det(A)}\sum_{k=1}^{n} A_{kj}b_k\,, \qquad (j = 1, \dots, n)\,.$$

Summan i högerledet kan uppfattas som utvecklingen efter kolonn j av en determinant $\det(B_i)$, där B_j är den matris som uppstår ur A då kolonn j ersätts med $\mathbf{b}$:

$$B_j = \begin{pmatrix} a_{11} & a_{12} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_2 & \cdots & a_{2n} \\ - & - & - & - & - & - \\ a_{n1} & a_{n2} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}.$$

Vi får därmed *Cramers regel* för komponenterna av $\mathbf{x}$:

$$x_j = \frac{\det(B_j)}{\det(A)}\,, \qquad (j = 1, \dots, n)\,.$$

Anmärkning. Cramers regel är inte lämplig för numeriska kalkyler. För det första kan den användas bara om $\det(A) \neq 0$. För det andra är den oekonomisk, eftersom den kräver väsentligt fler räkneoperationer än Gausseliminering. I nedanstående tabell jämförs antalet räkneoperationer vid lösning av ekvationen $A\mathbf{x} = \mathbf{b}$ medelst tre metoder: (1) genom Gausseliminering, (2) genom användning av formeln $\mathbf{x} = A^{-1}\mathbf{b}$ (där A^{-1} räknas ut genom Gausseliminering), (3) genom Cramers regel. Som räkneoperationer betraktas här addition (= subtraktion) och multiplikation (= division).

Metod	Antal operationer
Gausseliminering	$\approx \frac{2}{3}n^3$
$\mathbf{x} = A^{-1}\mathbf{b}$	$\approx 2n^3$
Cramers regel	$\approx \frac{2}{3}n^4$

Om $n = 10$ så kräver alltså Cramers regel ungefär 10 gånger fler operationer än lösning genom Gausseliminering, om $n = 100$ kräver den ungefär 100 gånger fler operationer.

Övningsuppgifter

1. (*a*) Räkna ut determinanten för matriserna

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{och} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

(*b*) Beräkna determinanten för matrisen

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

genom att gausseliminera A till en uppåt triangulär form U.

2. En matris A är *skevsymmetrisk* om $A^T = -A$. Visa att för varje skevsymmetrisk n/n-matris A gäller att $\det(A) = 0$ om n är ett udda heltal.

3. Bestäm determinanterna för

(*a*) matrisprodukten $A = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 2 \end{pmatrix}$;

(*b*) den uppåt triangulära matrisen $U = \begin{pmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 2 \end{pmatrix}$;

(*c*) den nedåt triangulära matrisen U^T;

(*d*) inversen U^{-1};

(*e*) matrisen $M = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 1 & 2 & 2 \\ 4 & 4 & 8 & 8 \end{pmatrix}$.

4. (*a*) Visa att om Q är en ortogonal matris så gäller $\det(Q) = 1$ eller $\det(Q) = -1$ (se uppgift 20 i kapitel 3).

(*b*) För vilka 2/2-matriser A gäller att $\det(3A) = 3\det(A)$?

5. Beräkna determinanten

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}.$$

6. Räkna ut den n-radiga determinanten

$$\begin{vmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ - & - & - & - & - \\ 1 & 1 & 1 & \cdots & 0 \end{vmatrix}.$$

7. Visa att

$$\begin{vmatrix} a+b & c+a & b+c \\ c+a & a+b & b+c \\ b+c & a+b & c+a \end{vmatrix} = 2(a+b+c) \begin{vmatrix} 1 & c & b \\ 1 & b & b \\ 1 & b & a \end{vmatrix}.$$

8. Beräkna determinanterna

$$(a) \begin{vmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 5 \\ 1 & -1 & 2 & 3 \end{vmatrix}, \quad (b) \begin{vmatrix} 1 & -2 & 3 & 1 \\ 5 & -9 & 6 & 3 \\ -1 & 2 & -6 & -2 \\ 2 & 8 & 6 & 1 \end{vmatrix}, \quad (c) \begin{vmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix}.$$

9. Visa att
$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)^3 .$$

10. Visa, utan att utveckla determinanterna, att
$$\begin{vmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_1 & b_2 \\ c_1 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_1 \\ c_2 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix} .$$

11. Räkna ut determinanaten
$$\begin{vmatrix} 4711 & 4712 & 4713 \\ 4712 & 4713 & 4714 \\ 4713 & 4714 & 4716 \end{vmatrix} .$$

12. Använd determinantformeln för en matrisinvers till att beräkna A^{-1}, då
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} .$$

13. Lös följande system med hjälp av Cramers regel:
$$\begin{aligned} x + 4y - z &= 1 \\ x + y + z &= 0 \\ 2x + 3z &= 0 . \end{aligned}$$

14. Visa att matriselementen i A^{-1} är heltal om A är inverterbar, alla matriselement i A är heltal och $\det(A)$ är 1 eller -1.

15. Visa att
$$\begin{vmatrix} 1 & 0 & .. & a_1 \\ 0 & 1 & .. & a_2 \\ - & - & - & - \\ 0 & 0 & .. & a_n \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 & .. & b_1 \\ 0 & 1 & .. & b_2 \\ - & - & - & - \\ 0 & 0 & .. & b_n \end{vmatrix} = \begin{vmatrix} 1+a_1b_1 & a_1b_2 & ... & a_1b_n \\ a_2b_1 & 1+a_2b_2 & ... & a_2b_n \\ - & - & - & - \\ a_nb_1 & a_nb_2 & ... & a_nb_n \end{vmatrix} .$$

16. Visa att om i en matris A summan av elementen i varje rad är noll, så är $\det(A) = 0$.

17. Räkna ut $\det(A)$, då n/n-matrisen $A = (a_{ik})$ har matriselementen $a_{ik} = i + k$ för varje i och k i mängden $\{1, \ldots, n\}$.

18. Bestäm med hjälp av determintteori för vilka värden på k följande matriser inte är inverterbara:
$$(a) \quad A = \begin{pmatrix} k-2 & 1 \\ -5 & k+4 \end{pmatrix}, \qquad (b) \quad B = \begin{pmatrix} k-4 & 0 & 0 \\ 0 & k & 2 \\ 0 & 3 & k-1 \end{pmatrix} .$$

19. I en triangel med sidorna a, b och c må de motstående vinklarna vara α, β respektive γ. Med hjälp av elementär trigonometri fås sambanden
$$\begin{aligned} b\cos\gamma + c\cos\beta &= a, \\ c\cos\alpha + c\cos\gamma &= b, \\ a\cos\beta + c\cos\alpha &= c. \end{aligned}$$
Härled med hjälp av Cramers regel formler for vinklarnas cosiner.

7. Egenvärden och egenvektorer

Låt A beteckna en n/n-matris. I vissa riktningar $\mathbf{x} \neq \mathbf{0}$ beter sig matrisen A enkelt i den meningen att $\mathbf{x}$ och $A\mathbf{x}$ råkar vara parallella:

Definition 7.1. Talet λ sägs vara ett *egenvärde* till A om

$$A\mathbf{x} = \lambda\mathbf{x}$$

för någon vektor $\mathbf{x} \neq \mathbf{0}$. Vektorn $\mathbf{x}$ sägs då vara en *egenvektor* till A som svarar mot egenvärdet λ.

Observera att om $\mathbf{x}$ är en egenvektor till A som svarar mot ett visst egenvärde λ, så är för varje $c \neq 0$ också $c\mathbf{x}$ en egenvektor som svarar mot egenvärdet λ. Längden av vektorn $\mathbf{x}$ är alltså irrelevant, endast riktningen har någon betydelse.

Egenvärden och egenvektorer är karakteristika för en matris A och berättar en hel del om matrisens egenskaper precis som t.ex. typen och rangen också gör det. Vi kommer att se att i många fall innehåller egenvärdena och egenvektorerna t.o.m. **all** information om A.

Exempel 7.1. För en enhetsmatris I gäller att $I\mathbf{x} = \mathbf{x}$ för varje $\mathbf{x}$. Således är talet 1 ett egenvärde till I och varje $\mathbf{x} \neq 0$ är en motsvarande egenvektor.

Om λ är ett egenvärde till A och $\mathbf{x}$ är en motsvarande egenvektor, så är

$$\begin{aligned} \left\{ \begin{matrix} A\mathbf{x} = \lambda\mathbf{x} \\ \text{för ngt } \mathbf{x} \neq 0 \end{matrix} \right\} &\iff \left\{ \begin{matrix} (A - \lambda I)\mathbf{x} = 0 \\ \text{för ngt } \mathbf{x} \neq 0 \end{matrix} \right\} \\ &\iff A - \lambda I \quad \text{är singulär} \\ &\iff \det(A - \lambda I) = 0\,. \end{aligned}$$

Således gäller:

Sats 7.1. *Om A är en n/n-matris, så fås A:s egenvärden genom att man löser den s.k.* karakteristiska ekvationen *(också kallad* sekularekvationen*)*

$$\det(A - \lambda I) = 0\,.$$

Därefter fås de egenvektorer, som svarar mot ett visst egenvärde λ, genom att man löser den homogena ekvationen $(A - \lambda I)\mathbf{x} = 0$.

Då man utvecklar determinanten i den karakteristiska ekvationen, fås alltid en polynomekvation av n:te graden, i vilken koefficienten för λ^n är $(-1)^n$. En sådan har ju alltid exakt n (reella eller komplexa) lösningar då multipliciteten beaktas. Den karakteristiska ekvationen kan alltså skrivas i den ekvivalenta formen

$$(\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_p)^{n_p} = 0\,,$$

där λ_k $(k = 1, \ldots, p)$ betecknar A:s **olika** egenvärden och n_k är den algebraiska *multipliciteten* hos egenvärdet λ_k. Om multipliciteten hos ett egenvärde är 1, 2, 3 ... sägs detta vara *enkelt*, *dubbelt*, *tredubbelt* ....

Definition 7.2. *Egenrummet* $V(\lambda)$ till en n/n-matris A med egenvärdet λ är nollrummet $N(A - \lambda I)$, dvs. lösningsmängden till $A\mathbf{x} = \lambda\mathbf{x}$.

Egenrummet $V(\lambda)$ är alltså ett underrum av $\mathbf{R}^n$ och består av nollvektorn och alla egenvektorer till A som svarar mot egenvärdet λ. Eftersom $N(A - 0 \cdot I) = N(A)$, så **är 0 ett egenvärde till A om och endast om matrisen A är singulär**.

Exempel 7.2. Den karakteristiska ekvationen för matrisen

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

är

$$0 = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 .$$

Ur denna löser vi ut egenvärdena 3 och 1. De egenvektorer som svarar mot egenvärdet 3 fås nu genom att man löser ekvationen $(A - 3I)\mathbf{x} = 0$:

$$A - 3I = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} .$$

Lösningarna är alltså $\mathbf{x} = t\,(1 \quad 1)^T$ och varje sådan vektor utom nollvektorn är en egenvektor som svarar mot egenvärdet 3. På samma sätt löser vi $(A - I)\mathbf{x} = 0$ med hjälp av räkneschemat

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

och får fram de egenvektorer $\mathbf{x} = t\,(-1 \quad 1)^T$ $(t \neq 0)$ som svarar mot egenvärdet 1.

Om matrisen A är reell och ett egenvärde λ inte är reellt (dvs. äkta komplext), så måste åtminstone någon komponent av en motsvarande egenvektor $\mathbf{x}$ vara icke-reell för att likheten $A\mathbf{x} = \lambda\mathbf{x}$ skall kunna gälla. Om däremot både A och egenvärdet λ är reella, så kan motsvarande egenvektorer $\mathbf{x}$ väljas reella.

Exempel 7.3. Den karakteristiska ekvationen för matrisen

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

är $\lambda^2 + 1 = 0$, varför egenvärdena är $\pm i$. De egenvektorer som svarar mot t.ex. egenvärdet i har formen $t\,(1 \quad -i)^T$, där $t \neq 0$ är ett reellt (eller komplext) tal.

Eftersom reella matriser således kan ha icke-reella egenvärden, så inser vi att teorin för egenvärden och egenvektorer egentligen borde utformas för komplexa matriser. I fortsättningen väljer vi emellertid alltid som exempel bara sådana matriser, som har reella egenvärden. I många fall är egenvärdena t.o.m. alltid automatiskt reella, såsom följande sats visar:

Sats 7.2. *Antag att A är en* symmetrisk *reell n/n-matris. Då gäller:*

(i) *Alla egenvärden till A är reella;*

(ii) *Om λ och μ är* olika *egenvärden till A så är $V(\lambda) \perp V(\mu)$.*

Bevis. (i) Låt λ vara ett egenvärde och låt $\mathbf{x}$ ($\neq \mathbf{0}$) vara en motsvarande egenvektor. Här måste vi temporärt acceptera att $\mathbf{x}$ kan ha icke-reella komponenter eftersom vi ännu inte vet att λ i själva verket måste vara reellt. Om $\mathbf{x} = (\,x_1 \;\; \ldots \;\; x_n\,)^T$, sätt $\mathbf{x}^* = (\,\overline{x_1} \;\; \ldots \;\; \overline{x_n}\,)$, där beteckningen $\overline{c} = a - ib$ står för konjugattalet till ett givet komplext tal $c = a + ib$. Om ekvationen $A\mathbf{x} = \lambda x$ multipliceras från vänster med $\mathbf{x}^*$ fås $\mathbf{x}^* A\mathbf{x} = \lambda \mathbf{x}^*\mathbf{x}$, dvs.

$$\lambda = \frac{\mathbf{x}^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}}\,.$$

Nämnaren $\mathbf{x}^*\mathbf{x} = |x_1|^2 + \cdots + |x_n|^2 \neq 0$ är en summa av kvadrater av belopp av komplexa tal och är således reell. Täljaren är ett komplext tal som sammanfaller med sitt eget konjugattal, ty eftersom $A = (a_{ik})$ är reell och symmetrisk, är

$$\begin{aligned}\overline{\mathbf{x}^* A\mathbf{x}} &= \overline{\sum_{i,k} \overline{x_i} a_{ik} x_k} = \sum_{i,k} x_i a_{ik} \overline{x_k} \\ &= \sum_{i,k} \overline{x_k} a_{ki} x_i = \mathbf{x}^* A\mathbf{x}\,.\end{aligned}$$

Då därmed både täljaren och nämnaren i uttrycket för λ är reella, så är λ reellt.

(ii) Vi skall visa att om vi tar godtyckliga (reella) vektorer $\mathbf{x} \in V(\lambda)$ och $\mathbf{y} \in V(\mu)$, så är $\mathbf{x} \perp \mathbf{y}$:

$$\begin{aligned}\left\{\begin{array}{l} A\mathbf{x} = \lambda\mathbf{x} \\ A\mathbf{y} = \mu\mathbf{y} \end{array}\right\} &\Longrightarrow \left\{\begin{array}{l} \mathbf{y}^T A\mathbf{x} = \lambda\mathbf{y}^T\mathbf{x} \\ \mathbf{x}^T A\mathbf{y} = \mu\mathbf{x}^T\mathbf{y} \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \mathbf{y}^T A\mathbf{x} = \lambda\mathbf{y}^T\mathbf{x} \\ \mathbf{y}^T A\mathbf{x} = \mu\mathbf{y}^T\mathbf{x} \end{array}\right\} \\ &\Longrightarrow \quad (\lambda - \mu)\mathbf{y}^T\mathbf{x} = 0 \quad \Longrightarrow \quad \mathbf{x} \perp \mathbf{y}\,. \diamondsuit\end{aligned}$$

För en matris A som inte är symmetrisk, behöver det inte gälla att egenvektorer som svarar mot olika egenvärden är ortogonala. Däremot måste de nog vara linjärt oberoende:

Sats 7.3. *Antag att $\lambda_1, \ldots, \lambda_p$ är* olika *egenvärden till en n/n-matris och låt $\mathbf{x}_1, \ldots, \mathbf{x}_p$ vara en uppsättning motsvarande egenvektorer, så att*

$$A\mathbf{x}_k = \lambda_k \mathbf{x}_k \qquad (k = 1, \ldots, p)\,.$$

Då är $\mathbf{x}_1, \ldots, \mathbf{x}_p$ linjärt oberoende.

Bevis. Satsen bevisas enklast genom induktion. Vi konstaterar först att påståendet gäller om $p = 1$. Sedan antar vi att påståendet gäller om vi har $p = k - 1$ egenvektorer $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}$ och skall bevisa att påståendet också gäller då antalet är $p = k$. Vi bildar därför en linjärkombination av de k egenvektorerna och sättar denna lika med $\mathbf{0}$:

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k = \mathbf{0}\,. \tag{1}$$

Vi multiplicera från vänster med A, beaktar att $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$ och får:

$$(2) \qquad c_1\lambda_1\mathbf{x}_1 + \cdots + \mathbf{c}_k\lambda_k\mathbf{x}_k = \mathbf{0}\,.$$

Vektorn $\mathbf{x}_k$ elimineras genom att vi bildar skillnaden mellan (2) och λ_k gånger (1):

$$c_1(\lambda_1 - \lambda_k)\mathbf{x}_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}\,.$$

Nu följer att $c_1 = \cdots = c_{k-1} = 0$, eftersom $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}$ är linjärt oberoende. Enligt (1) är därför $c_k\mathbf{x}_k = \mathbf{0}$ och därmed $c_k = 0$. Alltså är $\mathbf{x}_1,\ldots, \mathbf{x}_k$ linjärt oberoende. Induktionen ger att satsens påstående gäller för varje antal p av egenvektorer. $\diamondsuit$

Exempel 7.4. Matrisen

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

har den karakteristiska ekvationen

$$\begin{aligned} 0 &= \begin{vmatrix} 4-\lambda & 2 & 2 \\ 2 & 4-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 2-\lambda & 0 & -2+\lambda \\ 0 & 2-\lambda & -2+\lambda \\ 2 & 2 & 4-\lambda \end{vmatrix} \\ &= (2-\lambda)^2 \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 2 & 2 & 4-\lambda \end{vmatrix} = (2-\lambda)^2 \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 2 & 6-\lambda \end{vmatrix} \\ &= (2-\lambda)^2 \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 8-\lambda \end{vmatrix} = (2-\lambda)^2(8-\lambda)\,. \end{aligned}$$

Matrisen A har alltså egenvärdena 2 (dubbelt) och 8 (enkelt).

$\underline{\lambda = 2}$: En bas i egenrummet $V(2) = N(A - 2I)$ fås ur

$$A - 2I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

och resulterer i $\{(-1 \quad 1 \quad 0)^T, (-1 \quad 0 \quad 1)^T\}$.

$\underline{\lambda = 8}$: En motsvarande kalkyl

$$A - 8I = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

ger basen $\{(1 \quad 1 \quad 1)^T\}$ i $V(8) = N(A - 8I)$. Observera att vektorn $(1 \quad 1 \quad 1)^T$ är ortogonal mot basvektorerna i $V(2)$ men att basvektorerna i $V(2)$ inte (automatiskt) behöver bli ortogonala mot varandra (jfr. Sats 7.2).

Exempel 7.5. Efter en stunds kalkylerande finner man att matrisen

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

har den karakteristiska ekvationen $(\lambda - 1)(\lambda - 2)^2 = 0$.

$\underline{\lambda = 2}$: En bas i $V(2)$ fås ur

$$A - 2I = \begin{pmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

dvs. $\{(0 \quad 1 \quad 0)^T, (-1 \quad 0 \quad 1)^T\}$.

$\underline{\lambda = 1}$: Räkneschemat blir

$$A - I = \begin{pmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

En bas i $V(1)$ består alltså av en enda vektor $(-2 \quad 1 \quad 1)^T$, som uppenbarligen **inte** är ortogonal mot basvektorerna i $V(2)$. Däremot bildar unionen av baserna i $V(2)$ och $V(1)$ en linjärt oberoende mängd (jfr. Sats 7.3).

På basen av de två senaste exemplen kunde man få uppfattningen att om multipliciteten hos ett egenvärde λ är t.ex. 2, så innehåller en bas i $V(\lambda)$ precis 2 vektorer. Detta behöver inte alltid vara fallet men för symmetriska matriser är det så:

Anmärkning. Man kan visa att om A är en symmetrisk n/n-matris med den karakteristiska ekvationen

$$(\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_p)^{n_p} = 0,$$

där $\lambda_1, \ldots, \lambda_p$ är olika egenvärden och $\sum_i n_i = n$, så är

$$\dim V(\lambda_i) = n_i .$$

Vi ger ett exempel på en (icke-symmetrisk) matris, som saknar denna egenskap:

Exempel 7.6. Matrisen

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

har den karakteristiska ekvationen $0 = \det(A - \lambda I) = \lambda^2$, varför det enda egenvärdet 0 har multipliciteten 2. Matrisen är redan i reducerad echelonform, så vi kan direkt avläsa att $\{(1 \quad 0)\}$ är en bas i motsvarande egenrum $V(0)$. Dimensionen hos $V(0)$ är således bara 1, inte 2.

Anmärkning. Matriserna i exemplen i denna kurs är så små att det fortfarande är möjligt att utveckla determinanten i den karakteristiska ekvationen. Om matrisen är stor, blir detta oftast en uppgift som överstiger krafterna. Andra metoder krävs då för att man skall få fram egenvärdena till matrisen. Som ett exempel på en sådan metod nämner vi att koefficienterna i den karakteristiska ekvationen

$$\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 = 0$$

till en n/n-matris kan fås genom att man löser ekvationen

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ S_1 & 2 & 0 & \cdots & 0 & 0 \\ S_2 & S_1 & 3 & \cdots & 0 & 0 \\ - & - & - & - & - & - \\ S_{n-1} & S_{n-2} & S_{n-3} & \cdots & S_1 & n \end{pmatrix} \begin{pmatrix} c_{n-1} \\ c_{n-2} \\ c_{n-3} \\ - \\ c_0 \end{pmatrix} = - \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ - \\ S_n \end{pmatrix},$$

där $S_k = \operatorname{tr}(A^k)$ för $k = 1, 2, \ldots, n$ är spåret av potensen A^k av matrisen A. Egenvärdena kan sedan bestämmas genom att man löser den karakteristiska ekvationen med numeriska metoder.

Övningsuppgifter

1. Finn alla egenvärden och motsvarande egenvektorer till matriserna

$$(a)\quad \begin{pmatrix} 2 & 3 \\ 7 & 6 \end{pmatrix}, \qquad (b)\quad \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix}, \qquad (c)\quad \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

2. Bestäm egenvärden och baser i egenrummen till matriserna

$$(a)\quad \begin{pmatrix} 2 & 6 \\ 10 & 6 \end{pmatrix}, \qquad (b)\quad \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix},$$

$$(c)\quad \begin{pmatrix} 3 & 0 & 0 \\ -2 & 3 & -2 \\ 2 & 0 & 5 \end{pmatrix}, \qquad (d)\quad \begin{pmatrix} 5 & -12 & 6 \\ -1 & 5 & -1 \\ -5 & 18 & -6 \end{pmatrix}.$$

3. Bestäm egenvärden och baser i egenrummen för A^{62} då A är matrisen

$$(a)\quad \begin{pmatrix} 1 & -1 & -1 \\ -2 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \qquad (b)\quad \begin{pmatrix} 2 & -4 & 4 \\ -2 & 0 & 4 \\ -2 & 2 & 2 \end{pmatrix}.$$

4. Visa att den karakteristiska ekvationen för en (reell) 2/2-matris A kan skrivas

$$\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = 0\,.$$

   Bestäm villkoret för att A skall ha enbart reella egenvärden samt verifiera att detta villkor är uppfyllt för alla symmetriska 2/2-matriser.

5. Visa att A och A^T har samma egenvärden. Hur är två egenvektorer $\mathbf{x}$ och $\mathbf{y}$ till A respektive A^T som svarar mot olika egenvärden relaterade till varandra?

6. Visa att om A är en uppåt (nedåt) triangulär matris eller en diagonalmatris, så är A:s egenvärden precis matriselementen i huvuddiagonalen i A.

7. En matris A sägs vara *nilpotent* om $A^k = 0$ för något positivt heltal k. Visa att en nilpotent matris bara har egenvärdet 0.

8. Låt $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ vara ett polynom och sätt $p(A) = a_0 I + a_1 A + \cdots + a_n A^n$. Visa att om λ är ett egenvärde till A, så är $p(\lambda)$ ett egenvärde till $p(A)$.

8. Ortogonala projektioner

Antag att $\mathbf{a}$ $(\neq 0)$ och $\mathbf{b}$ är två vektorer i $\mathbf{R}^n$. Vi skall bilda den ortogonala projektionen av $\mathbf{b}$ på det endimensionella underrummet $L = \operatorname{spn}\{\mathbf{a}\}$. Enligt resonemanget i beviset av Schwarz olikhet är $\|\mathbf{b} - t\mathbf{a}\|^2$ minimalt då t har värdet $t_0 = \mathbf{a}^T\mathbf{b}/\|\mathbf{a}\|^2$.

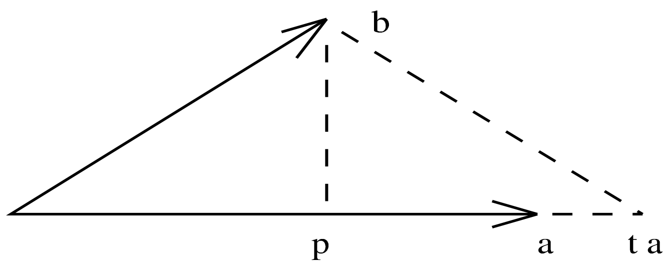

fig. 1

Om vi sätter $\mathbf{p} = t_0\mathbf{a}$, är $\mathbf{b} - \mathbf{p}$ ortogonalt mot L så att $\mathbf{b}$ kan skrivas som en summa $\mathbf{b} = \mathbf{p} + (\mathbf{b} - \mathbf{p})$, där $\mathbf{p} \in L$ och $\mathbf{b} - \mathbf{p} \in L^\perp$. Vektorn

(1) $$\mathbf{p} = t_0\mathbf{a} = \frac{\mathbf{a}^T\mathbf{b}}{\|\mathbf{a}\|^2}\mathbf{a}\,,$$

som vi också kan skriva

(2) $$\mathbf{p} = \frac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2}\mathbf{b}\,,$$

är alltså (den ortogonala) projektionen av $\mathbf{b}$ på L eller enklare *projektionen av* $\mathbf{b}$ *på* $\mathbf{a}$. Vi vet redan att $\mathbf{p}$ och $\mathbf{b} - \mathbf{p}$ är ortogonala men detta kan också verifieras direkt med hjälp av (1):

$$\mathbf{p}^T(\mathbf{b} - \mathbf{p}) = \mathbf{p}^T\mathbf{b} - \mathbf{p}^T\mathbf{p} = \frac{(\mathbf{a}^T\mathbf{b})^2}{\|\mathbf{a}\|^2} - \left(\frac{\mathbf{a}^T\mathbf{b}}{\|\mathbf{a}\|^2}\right)^2 \|\mathbf{a}\|^2 = 0\,.$$

På grund av denna ortogonalitet gäller (jfr. Pytagoras teorem):

$$\begin{aligned}\|\mathbf{b}\|^2 &= \mathbf{b}^T\mathbf{b} = (\mathbf{p} + (\mathbf{b} - \mathbf{p}))^T(\mathbf{p} + (\mathbf{b} - \mathbf{p})) \\ &= \mathbf{p}^T\mathbf{p} + 2\mathbf{p}^T(\mathbf{b} - \mathbf{p}) + (\mathbf{b} - \mathbf{p})^T(\mathbf{b} - \mathbf{p}) = \|\mathbf{p}\|^2 + \|\mathbf{b} - \mathbf{p}\|^2\,.\end{aligned}$$

Om $\mathbf{a}$ är en enhetsvektor $\mathbf{v}$, antar (1) den enklare formen

$$\mathbf{p} = (\mathbf{v}^T\mathbf{b})\mathbf{v} = (\|\mathbf{b}\| \cos\varphi)\mathbf{v}\,,$$

där φ är vinkeln mellan $\mathbf{b}$ och $\mathbf{a}$.

Exempel 8.1. För att projicera $\mathbf{b} = (3 \quad 1 \quad 2)^T$ på $\mathbf{a} = (1 \quad 1 \quad 1)^T$, räknar vi först ut enhetsvektorn

$$\mathbf{v} = \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{1}{\sqrt{3}}(1 \quad 1 \quad 1)^T .$$

Projektionen blir då

$$\mathbf{p} = (\mathbf{v}^T\mathbf{b})\mathbf{v} = \left(\frac{1}{\sqrt{3}} \cdot 6\right) \frac{1}{\sqrt{3}}(1 \quad 1 \quad 1)^T = 2(1 \quad 1 \quad 1)^T .$$

I nästa avsnitt skall vi generalisera formel (2) till det fall att en vektor **b** projiceras på ett underrum med godtycklig dimension. Först behöver vi emellertid några förberedande resultat:

Sats 8.1. *För en m/n-matrix A gäller:*

$$\begin{array}{ll} (i) & A^TA \text{ är en symmetrisk } n/n\text{-matris;} \\ (ii) & A^TA \text{ och } A \text{ har samma rang.} \end{array}$$

Bevis. (i) Matrisen A^TA är en n/n-matris. Att den är symmetrisk följer ur

$$(A^TA)^T = A^TA^{TT} = A^TA .$$

(ii) Vi visar först att A^TA och A har samma nollrum: Enligt Sats 5.7 (ii) gäller att $N(A^TA) \supseteq N(A)$. Den omvända inklusionen följer ur

$$\begin{aligned} \mathbf{x} \in N(A^TA) \quad &\Rightarrow \quad A^TA\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x}^TA^TA\mathbf{x} = 0 \quad \Rightarrow \quad (A\mathbf{x})^TA\mathbf{x} = 0 \\ &\Rightarrow \quad \|A\mathbf{x}\|^2 = 0 \quad \Rightarrow \quad A\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} \in N(A) . \end{aligned}$$

Alltså är $N(A^TA) = N(A)$. Enligt Sats 4.14 är nu

$$r(A^TA) = n - \dim N(A^TA) = n - \dim N(A) = r(A) . \diamondsuit$$

En följd av Sats 8.1 (ii) är:

Korollarium 8.1.1. *Om A är en m/n-matris med rangen n så är A^TA inverterbar.*

Bevis. Eftersom A^TA är en n/n-matris med rangen n, är A^TA icke-singulär och därmed inverterbar (Sats 3.9). $\diamondsuit$

Exempel 8.2. I matrisen

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

är kolonnerna linjärt oberoende. Därför är den symmetriska matrisen

$$A^TA = \begin{pmatrix} 14 & 7 \\ 7 & 17 \end{pmatrix}$$

inverterbar (vilket också lätt verifieras direkt).

Inkonsistenta ekvationer och projektionsmatriser

Antag att A är en m/n-matris och betrakta en ekvation $A\mathbf{x} = \mathbf{b}$, där $\mathbf{b}$ är en godtycklig kolonnvektor i $\mathbf{R}^m$. Om ekvationen är inkonsistent så har den naturligtvis ingen lösning men i detta fall kan man i stället försöka bestämma ett sådant $\mathbf{x}$ att $A\mathbf{x}$ ligger så nära högerledet $\mathbf{b}$ som möjligt:

Definition 8.1. En vektor $\mathbf{x}$, som gör talet $\|A\mathbf{x} - \mathbf{b}\|^2$ så litet som möjligt, kallas en *minstakvadratlösning* till ekvationen $A\mathbf{x} = \mathbf{b}$.

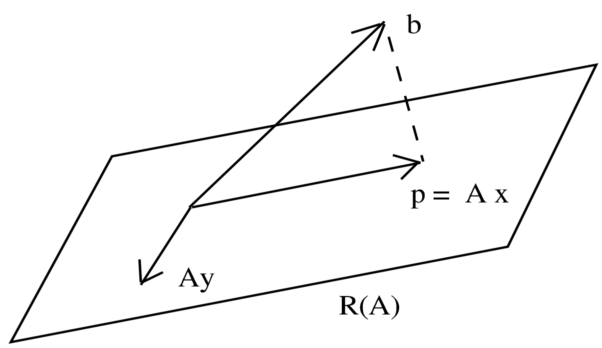

fig. 2

Talet $\|A\mathbf{x} - \mathbf{b}\|^2$ är ju kvadraten på avståndet från $\mathbf{b}$ till en variabel vektor $A\mathbf{x}$ i kolonnrummet $R(A)$. Då detta tal är så litet som möjligt, kommer $A\mathbf{x}$ således att vara den ortogonala projektionen $\mathbf{p}$ av $\mathbf{b}$ på $R(A)$.

Sats 8.2. *Minstakvadratlösningarna till ekvationen $A\mathbf{x} = \mathbf{b}$ är precis lösningarna till*

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Om $A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n)$ och kolonnerna $\mathbf{a}_i$ i A är linjärt oberoende, är A^TA inverterbar, varför minstakvadratlösningen till $A\mathbf{x} = \mathbf{b}$ i detta fall kan skrivas $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$. Den ortogonala projektionen av $\mathbf{b}$ på $R(A)$ blir alltså

$$\mathbf{p} = A\mathbf{x} = A(A^TA)^{-1}A^T\mathbf{b}.$$

Bevis. På grund av att $R(A) = \{A\mathbf{y} \mid \mathbf{y} \in \mathbf{R}^n\}$ fås följande kedja av ekvivalenta påståenden:

$$\begin{aligned}
\|A\mathbf{x} - \mathbf{b}\|^2 \text{ är minimalt} \quad &\Leftrightarrow \quad A\mathbf{x} - \mathbf{b} \perp A\mathbf{y} \text{ för varje } \mathbf{y} \\
&\Leftrightarrow \quad (A\mathbf{y})^T(A\mathbf{x} - \mathbf{b}) = 0 \text{ för varje } \mathbf{y} \\
&\Leftrightarrow \quad \mathbf{y}^T(A^TA\mathbf{x} - A^T\mathbf{b}) = 0 \text{ för varje } \mathbf{y} \\
&\Leftrightarrow \quad A^TA\mathbf{x} = A^T\mathbf{b}.
\end{aligned}$$

Minstakvadratlösningarna till $A\mathbf{x} = \mathbf{b}$ fås alltså genom att man löser ekvationen $A^TA\mathbf{x} = A^T\mathbf{b}$. Om $\mathbf{a}_1, \ldots, \mathbf{a}_n$ är linjärt oberoende så är enligt Korollarium 8.1.1

n/n-matrisen A^TA inverterbar. Då är $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$ den enda minstakvadratlösningen. För projektionen $\mathbf{p} = A\mathbf{x}$ av $\mathbf{b}$ fås då den formel som anges i satsen. $\diamondsuit$

Anmärkning. Observera att oändligt många vektorer $\mathbf{x}$ ibland kan ge upphov till samma minimala värde på $\|A\mathbf{x} - \mathbf{b}\|$. Detta inträffar varje gång ekvationen $A\mathbf{x} = \mathbf{0}$ har icke-triviala lösninger, eftersom A^TA och A har samma nollrum enligt Sats 8.1.

Låt oss se vad som händer om $\mathbf{b} \in R(A)$. Då är $\|A\mathbf{x} - \mathbf{b}\|^2$ minimalt, dvs. noll, precis då $\mathbf{p} = A\mathbf{x} = \mathbf{b}$. Minstakvadratlösningarna till $A\mathbf{x} = \mathbf{b}$ kommer alltså i detta fall att vara precis lösningarna till $A\mathbf{x} = \mathbf{b}$.

Ett annat extremt fall har vi om $\mathbf{b}$ är ortogonal mot $R(A)$. Då är $\mathbf{b}$ ortogonal mot kolonnerna i A så att $A^T\mathbf{b} = \mathbf{0}$. Minstakvadratlösningarna består då av alla $\mathbf{x} \in N(A^TA) = N(A)$ och projektionen av $\mathbf{b}$ på $R(A)$ blir $\mathbf{p} = A\mathbf{x} = \mathbf{0}$, vilket man kunde vänta sig.

Exempel 8.3. Räkneschemat för att bestämma minstakvadratlösningarna till ekvationen $A\mathbf{x} = \mathbf{b}$ är $(A^TA \mid A^T\mathbf{b})$. Detta räknar man enklast ut genom att bilda matrisprodukten $A^T(A \mid \mathbf{b})$. Om

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 2 & 1 & 2 \end{pmatrix} \quad \text{och} \quad \mathbf{b} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix},$$

ställer vi alltså upp

$$\begin{aligned} A^T(A \mid \mathbf{b}) &= \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 & | & 4 \\ 0 & 1 & 2 & | & 3 \\ 2 & 1 & 2 & | & 0 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 3 & 6 & | & 4 \\ 3 & 3 & 6 & | & 7 \\ 6 & 6 & 12 & | & 14 \end{pmatrix} \to \cdots \to \begin{pmatrix} 1 & 0 & 0 & | & -3/2 \\ 0 & 1 & 2 & | & 23/6 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}, \end{aligned}$$

vilker ger minstakvadratlösningarna

$$\mathbf{x} = \begin{pmatrix} -3/2 \\ 23/6 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} \qquad (t \in \mathbf{R}).$$

Antag nu att U är ett underrum av $\mathbf{R}^n$.

Definition 8.2. Den n/n-matris P, för vilken $P\mathbf{x}$ för varje $\mathbf{x} \in \mathbf{R}^n$ är (den ortogonala) projektionen av $\mathbf{x}$ på U, kallas *projektionsmatrisen* (för ortogonal projektion) på U.

Projektionsmatrisen P på U kan konstrueras på följande sätt: Om $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ är en bas i U, kan man låta A vara den matris som har $\mathbf{a}_1, \ldots, \mathbf{a}_k$ som kolonner. Då är ju $U = R(A)$ och matrisen

$$P = A(A^TA)^{-1}A^T \tag{3}$$

är enligt Sats 8.2 den sökta projektionsmatrisen på U.

Anmärkning. Om matrisen A innehåller bara en kolonn $\mathbf{a}$ så är $A^T A = \mathbf{a}^T \mathbf{a} = \|\mathbf{a}\|^2$ en 1/1-matris, dvs. en skalär. Dess invers är därför $1/\|\mathbf{a}\|^2$ och enligt (3) är $\mathbf{a}\mathbf{a}^T/\|\mathbf{a}\|^2$ projektionsmatrisen på $R(A) = \operatorname{spn}\{\mathbf{a}\}$. Detta visar att (2) är ett specialfall av (3).

Sats 8.3. (*a*) *Antag att P är projektionsmatrisen på ett underrum U av $\mathbf{R}^n$. Då har P egenskaperna*

$$(i) \quad P^2 = P\,,$$
$$(ii) \quad P^T = P\,.$$

Omvänt gäller att om en matris P har egenskaperna (i) och (ii), så är P projektionsmatrisen på underrummet $U = R(P)$.

(*b*) *$I - P$ är projektionsmatrisen på $U^\perp = N(P)$.*

Bevis. (*a*) Tag en bas i U och bilda en matris A som innehåller basvektorerna som kolonner. Då kan P skrivas som i (3). Alltså är

$$P^2 = A(A^TA)^{-1}A^T\, A(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = P\,,$$

eftersom $A^TA(A^TA)^{-1} = I$. Vidare är enligt Sats 3.10

$$\begin{aligned} P^T = (A(A^TA)^{-1}A^T)^T &= A^{TT}((A^TA)^{-1})^TA^T \\ &= A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T = P\,, \end{aligned}$$

så att både (*i*) och (*ii*) gäller. Om vi omvänt antar att P har egenskaperna (*i*) och (*ii*) så gäller för ett godtyckligt men fixerat $\mathbf{x}$ att

$$(P\mathbf{y}, \mathbf{x} - P\mathbf{x}) = (P\mathbf{y})^T(\mathbf{x} - P\mathbf{x}) = \mathbf{y}^TP^T\mathbf{x} - \mathbf{y}^TP^2\mathbf{x} = \mathbf{y}^TP\mathbf{x} - \mathbf{y}^TP\mathbf{x} = 0$$

för varje $\mathbf{y}$. Alltså är $\mathbf{x} - P\mathbf{x}$ ortogonal mot $R(P)$, dvs. $P\mathbf{x}$ är projektionen av $\mathbf{x}$ på underrummet $R(P)$.

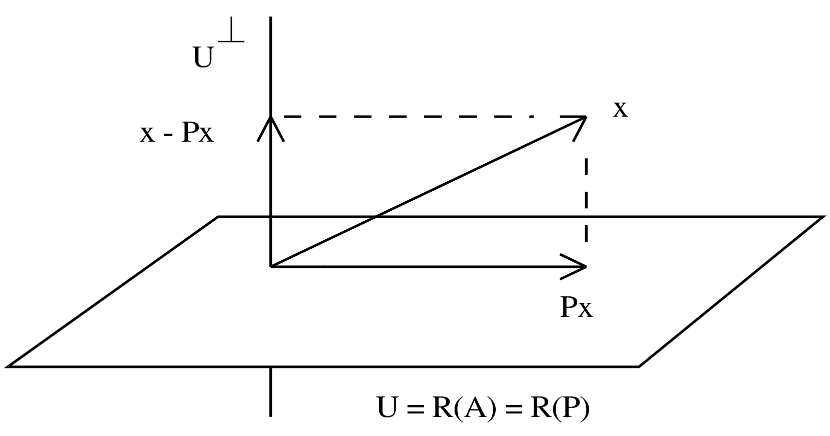

fig. 3

(*b*) Vi har just visat att vektorn $(I - P)\mathbf{x} = \mathbf{x} - P\mathbf{x}$ är ortogonal mot $U = R(A) = R(P)$ för varje $\mathbf{x}$. Alltså är $(I-P)\mathbf{x} \in U^\perp$. Varje $\mathbf{x}$ kan alltså skrivas som en summa $\mathbf{x} =$

$P\mathbf{x}+(I-P)\mathbf{x}$, där $P\mathbf{x} \in U$ och $(I-P)\mathbf{x} \in U^\perp$, vilket ger att $I-P$ är projektionsmatrisen på underrummet $U^\perp$. $\diamondsuit$

Exempel 8.4. Om man skall räkna ut projektionsmatrisen på det plan U i $\mathbf{R}^3$, som har ekvationen $x+3y+z=0$ i $\mathbf{R}^3$, kan man först bestämma en bas $\{\mathbf{a}_1, \mathbf{a}_2\}$ i planet, sätta $A = (\,\mathbf{a}_1 \quad \mathbf{a}_2\,)$ och sedan använda (3). Det är emellertid enklare att först bestämma projektionsmatrisen P' på det **lågdimensionella** $U^\perp = \operatorname{spn}\{\mathbf{a}\}$, där $\mathbf{a} = (\,1 \quad 3 \quad 1\,)^T$ är normalvektorn till U, och sedan räkna ut $P = I - P'$ (se Sats 8.3). I detta fall innehåller A bara kolonnen $\mathbf{a}$. Enligt (3) är

$$P' = A(A^TA)^{-1}A^T = \frac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2} = \frac{1}{11}\begin{pmatrix}1\\3\\1\end{pmatrix}(\,1 \quad 3 \quad 1\,) = \frac{1}{11}\begin{pmatrix}1&3&1\\3&9&3\\1&3&1\end{pmatrix},$$

och

$$P = I - P' = \frac{1}{11}\begin{pmatrix}11&&\\&11&\\&&11\end{pmatrix} - \frac{1}{11}\begin{pmatrix}1&3&1\\3&9&3\\1&3&1\end{pmatrix} = \frac{1}{11}\begin{pmatrix}10&-3&-1\\-3&2&-3\\-1&-3&10\end{pmatrix}.$$

Spektralframställningen

En viktig användning av projektionsmatriser har vi i samband med egenvärden och egenvektorer:

Sats 8.4. *Antag att A är en symmetrisk n/n-matris. Låt $\lambda_1, \ldots, \lambda_p$ vara de* olika *egenvärdena för A. Låt vidare P_i beteckna projektionsmatrisen på egenrummet $V(\lambda_i)$ $(i = 1, \ldots, p)$. Då gäller den s.k.* spektralframställningen *av A:*

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_p P_p\,.$$

Bevis. Eftersom $\sum_i \dim V(\lambda_i) = n$ och egenrummen $V(\lambda_i)$ är parvis ortogonala, kan varje $\mathbf{x} \in \mathbf{R}^n$ skrivas

$$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_p\,,$$

där $\mathbf{x}_i = P_i\mathbf{x} \in V(\lambda_i)$ $(i = 1, \ldots, p)$. Genom att multiplicera från vänster med A fås

$$\begin{aligned} A\mathbf{x} &= A\mathbf{x}_1 + \cdots + A\mathbf{x}_p \\ &= \lambda_1\mathbf{x}_1 + \cdots + \lambda_p\mathbf{x}_p \\ &= \lambda_1 P_1\mathbf{x} + \cdots + \lambda_p P_p\mathbf{x} \\ &= (\lambda_1 P_1 + \cdots + \lambda_p P_p)\mathbf{x}\,. \end{aligned}$$

Eftersom detta gäller för varje $\mathbf{x}$, så är $A = \lambda_1 P_1 + \cdots + \lambda_p P_p$. $\diamondsuit$

Låt oss se hur spektralframställningen kan användas i praktiken. Eftersom A är symmetrisk, är $V(\lambda_i) \perp V(\lambda_k)$ om $i \neq k$. Följaktligen är

$$P_iP_k = 0 \quad \text{om } i \neq k\,,$$

ty för varje $\mathbf{x} \in \mathbf{R}^n$ är $P_k\mathbf{x} \in V(\lambda_k) \subseteq N(P_i)$, varför $P_i(P_k\mathbf{x}) = 0$ för varje $\mathbf{x}$. Om vi nu bildar potenser av A, t.ex. A:s kvadrat, försvinner alla korstermer:

$$\begin{aligned} A^2 &= (\lambda_1 P_1 + \cdots + \lambda_p P_p)(\lambda_1 P_1 + \cdots + \lambda_p P_p) \\ &= \lambda_1^2 P_1^2 + \cdots + \lambda_p^2 P_p^2 = \lambda_1^2 P_1 + \cdots + \lambda_p^2 P_p\,. \end{aligned}$$

Allmänt fås på samma sätt för en godtycklig heltalsexponent $m \geq 1$:

$$A^m = \lambda_1^m P_1 + \cdots + \lambda_p^m P_p\,. \tag{4}$$

Om $\lambda_i \neq 0$ för varje i så gäller denna formel också om exponenten är 0 eller ett negativt heltal, förutsatt att vi definierar A^0 och A^{-m} genom $A^0 = I$ och $A^{-m} = (A^{-1})^m$: För $m = 0$ har vi ju för varje $\mathbf{x}$ den uppdelning i en summa som vi har använt ovan,

$$I\mathbf{x} = \mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_p = P_1\mathbf{x} + \cdots + P_p\mathbf{x}\,.$$

Ur

$$(\lambda_1 P_1 + \cdots + \lambda_p P_p)\left(\frac{1}{\lambda_1}P_1 + \cdots + \frac{1}{\lambda_p}P_p\right) = P_1 + \cdots + P_p = I$$

följer att

$$A^{-1} = \frac{1}{\lambda_1}P_1 + \cdots + \frac{1}{\lambda_p}P_p\,.$$

Då bägge leden i denna formel upphöjs till potensen m, ser vi att (4) gäller också för negativa heltalsexponenter.

Exempel 8.5. En kalkyl visar att matrisen

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

har egenvärdena 3 och 1 och att vektorn $\mathbf{a}_1 = (\,1 \quad 1\,)^T$ respektive $\mathbf{a}_2 = (\,-1 \quad 1\,)^T$ bildar en bas i $V(3)$ respektive $V(1)$. Projektionsmatriserna på dessa egenrum är alltså

$$\begin{aligned} P_1 &= \frac{\mathbf{a}_1\mathbf{a}_1^T}{\|\mathbf{a}_1\|^2} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(\,1 \quad 1\,) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ P_2 &= \frac{\mathbf{a}_2\mathbf{a}_2^T}{\|\mathbf{a}_2\|^2} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}(\,1 \quad -1\,) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned}$$

och spektralframställningen av A blir $A = 3P_1 + P_2$. Heltalspotenser A^n av A fås nu genom att att man upphöjer egenvärdena till potensen n:

$$\begin{aligned} A^n &= 3^n P_1 + P_2 \\ &= 3^n \cdot \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 3^n + 1 & 3^n - 1 \\ 3^n - 1 & 3^n + 1 \end{pmatrix} \end{aligned}$$

för varje $n \in \mathbf{Z}$. Notera att detta kan användas t.ex. till att bestämma gränsvärdet av A^n då $n \to -\infty$: Eftersom gränsvärden av matriser bildas elementvis, är

$$\lim_{n \to -\infty} A^n = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = P_2\,.$$

Ortogonala matriser och projektionsmatriser

Vi har tidigare visat att om matrisen A har linjärt oberoende kolonner så är A^TA kvadratisk, symmetrisk och inverterbar. Låt oss nu göra det strängare antagandet att kolonnerna $\mathbf{a}_1, \ldots, \mathbf{a}_p$ i A bildar ett ON-system (se definition 5.5). I detta fall blir matrisen A^TA speciellt enkel:

$$(5) \qquad A^TA = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p \end{pmatrix} (\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_p) = \begin{pmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \ldots & \mathbf{a}_1^T\mathbf{a}_p \\ \vdots & & \vdots \\ \mathbf{a}_p^T\mathbf{a}_1 & \ldots & \mathbf{a}_p^T\mathbf{a}_p \end{pmatrix} = I_p \,.$$

För projektionsmatrisen $P = A(A^TA)^{-1}A^T$ på $R(A)$ får vi därför uttrycket

$$(6) \qquad P = AA^T \,.$$

Eftersom vi får minstakvadratlösningarna till $A\mathbf{x} = \mathbf{b}$ som lösningar till $A^TA\mathbf{x} = A^T\mathbf{b}$, är nu $\mathbf{x} = A^T\mathbf{b}$ den entydiga minstakvadratlösningen.

Formel (6) kan också skrivas med hjälp av beteckningarna för kolonnerna i A:

$$(7) \qquad P = AA^T = (\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_p) \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p \end{pmatrix} = \mathbf{a}_1\mathbf{a}_1^T + \cdots + \mathbf{a}_p\mathbf{a}_p^T \,.$$

Eftersom varje $\mathbf{a}_i$ är en enhetsvektor, dvs. ett ON-system bestående av en enda vektor, så är varje term $\mathbf{a}_i\mathbf{a}_i^T$ i summan projektionsmatrisen på $\operatorname{spn}\{\mathbf{a}_i\}$ enligt (6).

Definition 8.3. En kvadratisk matris Q är *ortogonal* om kolonnerna i Q bildar ett ON-system.

Ur (5) följer att en kvadratisk matris Q är ortogonal om och endast om $Q^TQ = I$. Detta är i sin tur ekvivalent med att Q^T är Q:s invers (Sats 4.17 (*ii*)). Vi har alltså följande karakterisering av en ortogonal matris:

Sats 8.5. *En kvadratisk matris Q är ortogonal om och endast om Q^{-1} existerar och $Q^{-1} = Q^T$.*

Enligt Sats 3.10 är $(Q^T)^{-1} = (Q^{-1})^T$ om Q^{-1} existerar. Om Q är ortogonal, är därför $(Q^T)^{-1} = (Q^T)^T$, dvs. också Q^T är en ortogonal matris:

Korollarium 8.5.1. *Om matrisen Q är ortogonal, är också Q^T ortogonal, dvs. också raderna i Q bildar ett ON-system.*

Ortogonala matriser har en **geometrisk** betydelse eller tolkning, som vi nu skall studera närmare:

Antag att Q är en ortogonal n/n-matris och låt $\mathbf{x}$ och $\mathbf{y}$ vara godtyckliga vektorer i $\mathbf{R}^n$. Mellan dessa har vi en viss vinkel u (om vektorerna är olika noll). Då man multiplicerar från vänster med Q erhålls två nya vektorer $Q\mathbf{x}$ och $Q\mathbf{y}$, mellan vilka vi har en vinkel v. För t.ex. $Q\mathbf{x}$ gäller

$$\|Q\mathbf{x}\|^2 = (Q\mathbf{x})^TQ\mathbf{x} = \mathbf{x}^TQ^TQ\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2 \,,$$

eftersom $Q^TQ = I$. Efter kvadratrotsutdragning fås

$$\|Q\mathbf{x}\| = \|\mathbf{x}\|.$$

Motsvarande formel gäller för vilken vektor som helst, t.ex. för $\mathbf{x}-\mathbf{y}$, så att

$$\|Q\mathbf{x} - Q\mathbf{y}\| = \|Q(\mathbf{x}-\mathbf{y})\| = \|\mathbf{x}-\mathbf{y}\|.$$

Denna formel säger att **avstånd bevaras** vid multiplikation från vänster med Q.

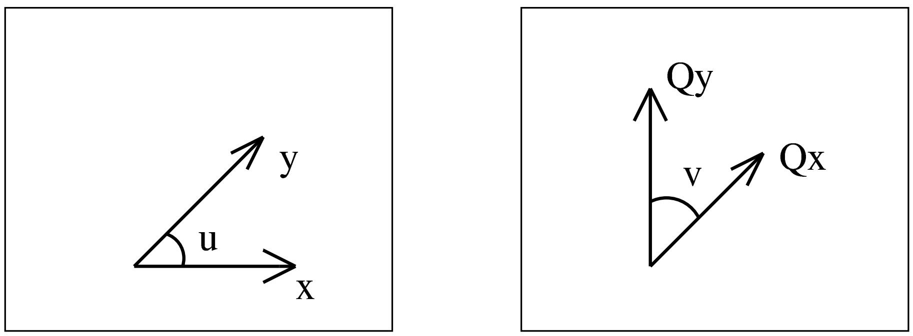

fig. 4

Om avstånd bevaras så **bevaras också vinklar**: En triangel förflyttas ju i $\mathbf{R}^n$ (se fig. 4) och då bevaras triangelns vinklar. Men man kan också direkt visa att $u = v$ med hjälp av att $Q^TQ = I$:

$$\cos v = \frac{(Q\mathbf{x})^T Q\mathbf{y}}{\|Q\mathbf{x}\|\,\|Q\mathbf{y}\|} = \frac{\mathbf{x}^T Q^T Q\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} = \cos u.$$

Sats 8.6. *Då vektorer i $\mathbf{R}^n$ multipliceras från vänster med en ortogonal n/n-matris bevaras både avstånd och vinklar.*

Exempel 8.6. Som ett exempel på en ortogonal matris betraktar vi

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

dvs. den matris som "roterar" xy-planet en vinkel θ moturs (se exempel 5.8). De två kolonnerna är ortogonala mot varandra och är enhetsvektorer (ty $\cos^2\theta + \sin^2\theta = 1$). Alltså är Q en ortogonal matris. Vi säger att Q är en *rotationsmatris.* Rotationsmatriser i $\mathbf{R}^3$ åstadkommer på motsvarande sätt en rotation med en viss vinkel kring en given rotationsaxel. Dessa matriser är också ortogonala men kan vara betydligt mera komplicerade.

Andra exempel på ortogonala matriser har vi i *speglingsmatriserna.* Låt V vara ett underrum i $\mathbf{R}^n$ och låt P vara projektionsmatrisen på V. För varje vektor $\mathbf{x}$ i $\mathbf{R}^n$ är vektorn

$$S\mathbf{x} = x + 2(P\mathbf{x} - \mathbf{x}) = 2P\mathbf{x} - \mathbf{x} = (2P - I)\mathbf{x}$$

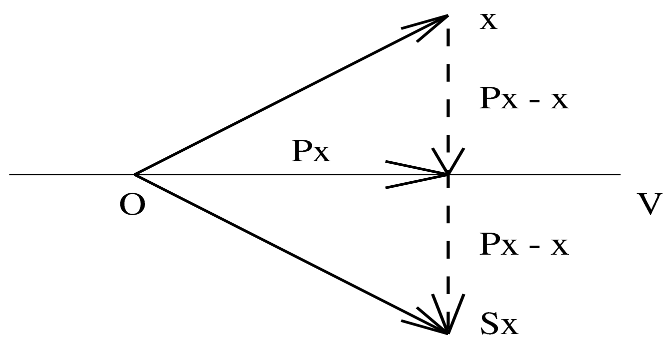

fig. 5

speglingen av $\mathbf{x}$ i V. Den matris som åstadkommer denna spegling, dvs.

$$S = 2P - I \tag{8}$$

kallas *speglingsmatrisen* i V. Uttryckt med hjälp av projektionsmatrisen $P' = I - P$ på $V^\perp$ har S formen

$$S = I - 2P'\,. \tag{9}$$

Eftersom $P^T = P = P^2$, är

$$\begin{aligned} S^T &= (2P-I)^T = 2P^T - I = 2P - I = S\,,\\ S^2 &= (2P-I)^2 = 4P^2 - 4P + I = I\,. \end{aligned}$$

Följaktligen är $S^TS = S^2 = I$, dvs. $S^{-1} = S^T$. Alltså är S en ortogonal matris.

Exempel 8.7. För att räkna ut speglingsmatrisen i planet V i $\mathbf{R}^3$ med ekvationen $2x - y + 3z = 0$ är det enklare att använda formel (9) än formel (8) på grund av att $V^\perp$ har **lägre dimension** än V. Planet V har normalvektorn $\mathbf{a} = (\,2 \quad -1 \quad 3\,)^T$. Således är

$$\begin{aligned} S &= I - 2P' = I - 2\frac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2}\\ &= \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} - 2\cdot\frac{1}{14}\begin{pmatrix} 2\\ -1\\ 3 \end{pmatrix}(\,2 \quad -1 \quad 3\,) = \frac{1}{7}\begin{pmatrix} 3 & 2 & -6\\ 2 & 6 & 3\\ -6 & 3 & -2 \end{pmatrix}. \end{aligned}$$

Anmärkning. Man kan visa att varje ortogonal n/n-matris är en produkt av högst n stycken speglingsmatriser i hyperplan i $\mathbf{R}^n$, dvs. i underrum med dimensionen $n-1$. Vidare kan man visa att en produkt av två sådana speglingsmatriser är en rotationsmatris kring en "axel" med dimensionen $n-2$. T.ex. de ortogonala 3/3-matriserna kan därför indelas i (i) speglingsmatriser i ett plan, (ii) rotationsmatriser och (iii) rotationsspeglingar. En rotationsspegling är en produkt av en rotationsmatris och en speglingsmatris för spegling i ett plan. Observera att varje permutationsmatris är en ortogonal matris: En enkel permutationsmatris är ju en speglingsmatris i ett hyperplan.

Gram–Schmidt-proceduren

Antag att ett antal linjärt oberoende vektorer $\mathbf{a}_1, \ldots, \mathbf{a}_p$ är givna. Vår uppgift är att konstruera **ortonormala** vektorer $\mathbf{q}_1, \ldots, \mathbf{q}_p$ (dvs. att *ortonormera*), sådana att

$$\operatorname{spn}\{\mathbf{q}_1, \ldots, \mathbf{q}_p\} = \operatorname{spn}\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}.$$

Med hjälp av *Gram–Schmidt-proceduren* utförs detta i två steg. I det första steget konstrueras ortogonala vektorer med samma spann som de ursprungliga (man *ortogonaliserar*). I det andra steget ändras längden av de ortogonala vektorerna så att de blir enhetsvektorer (man *normerar*).

Steg 1: Vi konstruerar stegvis nya vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_p$ på följande sätt: Sätt $\mathbf{v}_1 = \mathbf{a}_1$. Sätt $\mathbf{v}_2 = \mathbf{a}_2 - \lambda \mathbf{v}_1$ och välj λ så att $\mathbf{v}_1 \perp \mathbf{v}_2$. Vi kräver följaktligen att

$$0 = \mathbf{v}_1^T \mathbf{v}_2 = \mathbf{v}_1^T \mathbf{a}_2 - \lambda \|\mathbf{v}_1\|^2,$$

vilket ger att $\lambda = \mathbf{v}_1^T \mathbf{a}_2 / \|\mathbf{v}_1\|^2$. Alltså är

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{\mathbf{v}_1^T \mathbf{a}_2}{\|\mathbf{v}_1\|^2} \mathbf{v}_1.$$

Sätt sedan $\mathbf{v}_3 = \mathbf{a}_3 - \lambda_1 \mathbf{v}_1 - \lambda_2 \mathbf{v}_2$ och välj λ_1 och λ_2 så att $\mathbf{v}_1 \perp \mathbf{v}_3$ och $\mathbf{v}_2 \perp \mathbf{v}_3$. Eftersom $\mathbf{v}_1$ och $\mathbf{v}_2$ redan är ortogonala, antar dessa krav formen

$$\begin{aligned} 0 &= \mathbf{v}_1^T \mathbf{v}_3 = \mathbf{v}_1^T \mathbf{a}_3 - \lambda_1 \|\mathbf{v}_1\|^2, \\ 0 &= \mathbf{v}_2^T \mathbf{v}_3 = \mathbf{v}_2 \mathbf{a}_3 - \lambda_2 \|\mathbf{v}_2\|^2. \end{aligned}$$

Detta ger att $\lambda_1 = \mathbf{v}_1^T \mathbf{a}_3 / \|\mathbf{v}_1\|^2$ och $\lambda_2 = \mathbf{v}_2^T \mathbf{a}_3 / \|\mathbf{v}_2\|^2$, så att

$$\mathbf{v}_3 = \mathbf{a}_3 - \frac{\mathbf{v}_1^T \mathbf{a}_3}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\mathbf{v}_2^T \mathbf{a}_3}{\|\mathbf{v}_2\|^2} \mathbf{v}_2.$$

På detta sätt kan man fortsätta: Ansatsen för varje $\mathbf{v}_i$ är **den gamla vektorn $\mathbf{a}_i$ minus en linjärkombination av de tidigare definierade $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$**. Därefter väljs koefficienterna i linjärkombinationen så att $\mathbf{v}_i$ faktiskt blir ortogonal mot $\mathbf{v}_1, \ldots,$ $\mathbf{v}_{i-1}$. Som sista ekvation fås

$$\mathbf{v}_p = \mathbf{a}_p - \frac{\mathbf{v}_1^T \mathbf{a}_p}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \cdots - \frac{\mathbf{v}_{p-1}^T \mathbf{a}_p}{\|\mathbf{v}_{p-1}\|^2} \mathbf{v}_{p-1}.$$

Efter insättningar blir varje $\mathbf{v}_i$ en linjärkombination av vektorer $\mathbf{a}_i$ och omvänt kan varje $\mathbf{a}_i$ lösas ut som en linjärkombination av vektorer $\mathbf{v}_i$. Således är

$$\operatorname{spn}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} = \operatorname{spn}\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}.$$

Steg 2: Nu återstår bara att normera vektorerna: För varje $i = 1, \ldots, p$ sätter vi

$$\mathbf{q}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}, \quad \text{varvid} \quad \|\mathbf{q}_i\| = \left\| \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|} \right\| = \frac{1}{\|\mathbf{v}_i\|} \|\mathbf{v}_i\| = 1.$$

Vektorerna $\mathbf{q}_1, \dots, \mathbf{q}_p$ bildar nu ett ON-system med samma spann som de ursprungliga vektorerna $\mathbf{a}_1, \dots, \mathbf{a}_p$.

Gram–Schmidt-proceduren används ofta på egenvektorer till symmetriska matriser:

Exempel 8.8. Den symmetriska matrisen

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

visar sig ha den karakteristika ekvationen $(2-\lambda)^2(5-\lambda) = 0$. Efter uträkning finner vi i egenrummen $V(2)$ och $V(5)$ baser $\{\mathbf{a}_1, \mathbf{a}_2\}$ respektive $\{\mathbf{a}_3\}$, där

$$\mathbf{a}_1 = (-1 \quad 1 \quad 0)^T, \quad \mathbf{a}_2 = (-1 \quad 0 \quad 1)^T, \quad \mathbf{a}_3 = (1 \quad 1 \quad 1)^T .$$

Vektorn $\mathbf{a}_3$ är automatiskt ortogonal mot $\mathbf{a}_1$ och $\mathbf{a}_2$ (Sats 7.2 (ii)) men $\mathbf{a}_1$ och $\mathbf{a}_2$ är inte sinsemellan ortogonala. Vi tillämpar Gram–Schmidt-proceduren på dem, dvs. vi sätter $\mathbf{v}_1 = \mathbf{a}_1$ och $\mathbf{v}_2 = \mathbf{a}_2 - \lambda \mathbf{v}_1$ samt kräver att

$$0 = \mathbf{v}_1^T \mathbf{v}_2 = \mathbf{v}_1^T \mathbf{a}_2 - \lambda \|\mathbf{v}_1\|^2 = 1 - 2\lambda .$$

Ur detta följer att $\lambda = 1/2$. Alltså är

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{1}{2}\mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{2}\mathbf{w}_2 .$$

Vektorerna $\mathbf{v}_1$ och $\mathbf{v}_2$ skall sedan normeras i steg 2 (observera att $\mathbf{v}_2$ och $\mathbf{w}_2$har samma normering!). Samtidigt passar vi på att normera också $\mathbf{a}_3$:

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix},$$

$$\mathbf{q}_3 = \frac{\mathbf{a}_3}{\|\mathbf{a}_3\|} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

Vektorerna $\mathbf{q}_1$, $\mathbf{q}_2$ och $\mathbf{q}_3$ bildar nu ett ON-system av egenvektorer till A. Med hjälp av (7) kan vi lätt skriva ut spektralframställningen

$$A = 2(\mathbf{q}_1\mathbf{q}_1^T + \mathbf{q}_2\mathbf{q}_2^T) + 5\mathbf{q}_3\mathbf{q}_3^T .$$

Exempel 8.9. Om vektorerna

$$\mathbf{a}_1 = (1 \quad 1 \quad 0)^T, \quad \mathbf{a}_2 = (1 \quad 0 \quad 1)^T, \quad \mathbf{a}_3 = (0 \quad 1 \quad 1)^T$$

skall ortonormeras, sätter vi $\mathbf{v}_1 = \mathbf{a}_1 = (1 \quad 1 \quad 0)^T$ och

$$\mathbf{v}_2 = \mathbf{a}_2 - \lambda \mathbf{v}_1$$

samt kräver att

$$0 = \mathbf{v}_1^T \mathbf{v}_2 = \mathbf{v}_1^T \mathbf{a}_2 - \lambda \|\mathbf{v}_1\|^2 = 1 - 2\lambda\,.$$

Ur detta följer att $\lambda = 1/2$ och då är

$$\mathbf{v}_2 = \mathbf{a}_2 - \lambda \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{2} \mathbf{w}_2\,.$$

Sedan sätter vi $\mathbf{v}_3 = \mathbf{a}_3 - \lambda_1 \mathbf{v}_1 - \lambda_2 \mathbf{v}_2$ och kräver att

$$\begin{aligned} 0 &= \mathbf{v}_1^T \mathbf{v}_3 = \mathbf{v}_1^T \mathbf{a}_3 - \lambda_1 \|\mathbf{v}_1\|^2 = 1 - 2\lambda_1\,, \\ 0 &= \mathbf{v}_2^T \mathbf{v}_3 = \mathbf{v}_2^T \mathbf{a}_3 - \lambda_2 \|\mathbf{v}_2\|^2 = \frac{1}{2} - \frac{3}{2}\lambda_2\,. \end{aligned}$$

Detta ger att $\lambda_1 = 1/2$ och $\lambda_2 = 1/3$. Alltså är

$$\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{2}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \frac{2}{3} \mathbf{w}_3\,.$$

I steg 2 sätter vi till slut

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix},$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

QR-faktoriseringen

QR-faktoriseringen av en matris är i själva verket en skrivning i matrisform av Gram–Schmidt-ortonormeringen av matrisens kolonner:

Exempel 8.10. I föregående exempel är

$$\begin{aligned} \mathbf{a}_1 &= \mathbf{v}_1 & &= \sqrt{2}\mathbf{q}_1 \\ \mathbf{a}_2 &= \frac{1}{2}\mathbf{v}_1 + \mathbf{v}_2 & &= \frac{\sqrt{2}}{2}\mathbf{q}_1 + \frac{\sqrt{6}}{2}\mathbf{q}_2 \\ \mathbf{a}_3 &= \frac{1}{2}\mathbf{v}_1 + \frac{1}{3}\mathbf{v}_2 + \mathbf{v}_3 & &= \frac{\sqrt{2}}{2}\mathbf{q}_1 + \frac{\sqrt{6}}{6}\mathbf{q}_2 + \frac{2\sqrt{3}}{3}\mathbf{q}_3\,. \end{aligned}$$

Om vi sätter $A = (\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,)$ och $Q = (\,\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3\,)$, kan dessa likheter skrivas i matrisform

$$A = QR = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & \sqrt{6}/2 & \sqrt{6}/6 \\ 0 & 0 & 2\sqrt{3}/3 \end{pmatrix}.$$

Detta är QR-faktoriseringen av matrisen A.

Av exempel 8.10 inser man att varje matris A med linjärt oberoende kolonner kan faktoriseras på motsvarande sätt:

Sats 8.7. *Varje matris A med linjärt oberoende kolonner kan genom ortonormering av kolonnerna QR-faktoriseras,*

$$A = QR\,,$$

där kolonnerna i Q är de ortonormala vektorer, som är resultatet av Gram–Schmidt-proceduren, och R är en uppåt triangulär matris, som är inverterbar. Diagonalelementen i R är $r_{ii} = \|\mathbf{v}_i\|$, där $\mathbf{v}_i$ är de vektorer som definieras under ortogonaliseringsfasen. Om A är kvadratisk, är Q en ortogonal matris.

Volymer

Med hjälp av Gram–Schmidt-proceduren och QR-faktoriseringen skall vi härleda en formel för volymen V av den "parallellepiped" i $\mathbf{R}^n$ som spänns upp av n vektorer $\mathbf{a}_1$, $\ldots$, $\mathbf{a}_n$, dvs. den parallellepiped som har vektorerna $\mathbf{a}_i$ som "kanter".

Vi antar först att vektorerna $\mathbf{a}_i$ är parvis ortogonala. Då är $V = \|\mathbf{a}_1\| \cdots \|\mathbf{a}_n\|$. Om vi sätter $A = (\,\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_n\,)$ är å andra sidan

$$A^T A = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{pmatrix} (\,\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n\,) = \begin{pmatrix} \|\mathbf{a}_1\|^2 & & \\ & \ddots & \\ & & \|\mathbf{a}_n\|^2 \end{pmatrix},$$

vilket ger att

$$\det(A)^2 = \det(A^T) \cdot \det(A) = \det(A^T A) = \|\mathbf{a}_1\|^2 \cdots \|\mathbf{a}_n\|^2 = V^2\,.$$

Således är

$$V = |\det(A)|\,. \tag{10}$$

Vi skall visa att volymformeln (10) gäller generellt. Först konstaterar vi att den gäller om vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_n$ är linjärt beroende, eftersom både volymen V och determinanten då är noll. Vi kan därför i fortsättningen anta att $\mathbf{a}_1, \ldots, \mathbf{a}_n$ är linjärt oberoende men annars godtyckliga. Vi ortogonaliserar enligt Gram–Schmidt-proceduren och får ortogonala vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ samt erhåller en QR-faktorisering $A = QR$ av A. Av fig. 6 ser vi att det parallellogram, som $\mathbf{a}_1$ och $\mathbf{a}_2$ uppspänner i $\mathbf{R}^2$, har samma area som den rektangel, som $\mathbf{v}_1$ och $\mathbf{v}_2$ uppspänner. Motsvarande kommer att gälla för volymer i $\mathbf{R}^3$ och i högre dimensioner.

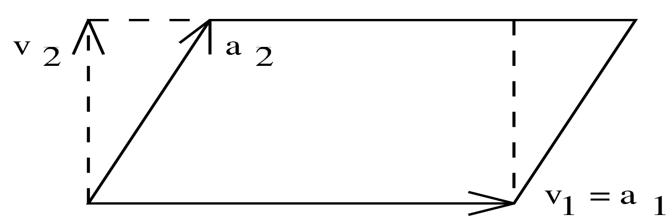

fig. 6

Med stöd av detta och räkneregel 8 för determinanter fås

$$\det(A) = \det(Q) \cdot \det(R) = \pm\|\mathbf{v}_1\| \cdots \|\mathbf{v}_n\| = \pm V\,.$$

Här har vi dessutom utnyttjat att diagonalelementen i R är $\|\mathbf{v}_i\|$ (Sats 8.7) samt att determinanten av en ortogonal matris är ± 1 (övningsuppgift 4 i kap. 6). Formel (10) är alltså giltig i varje tänkbart fall.

Övningsuppgifter

1. Konstruera en ortonormal bas i planet $x - y + z = 0$ i $\mathbf{R}^3$ samt bestäm projektionsmatrisen P på planet.
2. Bestäm på den räta linje som uppspänns av $\mathbf{a} = (1 \quad 1 \quad 1)^T$ den punkt $\mathbf{p}_1$, som ligger närmast punkten $\mathbf{b} = (2 \quad 4 \quad 4)^T$. Bestäm också den punkt $\mathbf{p}_2$ på spn$\{\mathbf{b}\}$, vilken ligger närmast $\mathbf{a}$.
3. (a) Bestäm minstakvadratlösningen till $A\mathbf{x} = \mathbf{b}$ genom att lösa $A^TA\mathbf{x} = A^T\mathbf{b}$, där
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}.$$
Bestäm projektionen $\mathbf{p}$ av $\mathbf{b}$ på $R(A)$.

   (b) och (c): Utför samma uppgift då A och $\mathbf{b}$ är följande matris respektive högerled:
$$(b) \quad \begin{pmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}; \qquad (c) \quad \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}.$$
4. Antag att P är projektionsmatrisen på en rät linje genom origo i xy-planet. Rita en figur för att beskriva effekten av speglingsmatrisen $H = I - 2P$. Förklara både geometriskt och algebraiskt varför $H^2 = I$.
5. (a) Projicera vektorn $\mathbf{b} = (0 \quad 3 \quad 0)^T$ på var och en av de ortonormala vektorerna $\mathbf{a}_1 = \frac{1}{3}(2 \quad 2 \quad -1)^T$ och $\mathbf{a}_2 = \frac{1}{3}(-1 \quad 2 \quad 2)^T$ och bestäm sedan dess projektion på planet spn$\{\mathbf{a}_1, \mathbf{a}_2\}$.

   (b) Bestäm också projektionen av $\mathbf{b} = (0 \quad 3 \quad 0)^T$ på $\mathbf{a}_3 = \frac{1}{3}(2 \quad -1 \quad 2)^T$, addera de tre projektionerna på endimensionella underrum och tolka resultatet. Varför är $P = \mathbf{a}_1\mathbf{a}_1^T + \mathbf{a}_2\mathbf{a}_2^T + \mathbf{a}_3\mathbf{a}_3^T = I$?
6. Bestäm en bas i $N(A)$ då
$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix}$$
och verifiera att $N(A) \perp R(A^T)$. Skriv vektorn $\mathbf{x} = (3 \quad 3 \quad 3)^T$ i formen $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, där $\mathbf{x}_1 \in R(A^T)$ och $\mathbf{x}_2 \in N(A)$. Vilka är koordinaterna för $\mathbf{x}_1$ i basen $\{(1 \quad 0 \quad 2)^T, (1 \quad 1 \quad 4)^T\}$ i $R(A^T)$?
7. Räkna ut spektralframställningen för matrisen A då
$$(a) \quad A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}, \qquad (b) \quad A = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$
Skriv ut en formel för A^n samt bestäm en tredjerot ur A, dvs. en matris X sådan att $X^3 = A$.

8. Visa att om Q_1 och Q_2 är ortogonala matriser så är också Q_1Q_2 en ortogonal matris.
9. Räkna ut en formel för avståndet från en punkt $\mathbf{y}$ i $\mathbf{R}^3$ till planet $\mathbf{n}^T\mathbf{x} = b$. Ledning: Antag temporärt att $\mathbf{x}_0$ är en punkt i planet och projicera $\mathbf{y} - \mathbf{x}_0$ på $\mathbf{n}$.
10. Ortonormera med hjälp av Gram–Schmidt-proceduren vektorerna

$$\begin{aligned}
&(a) \quad \mathbf{a}_1 = (\,0 \quad 0 \quad 1\,)^T, \quad \mathbf{a}_2 = (\,0 \quad 1 \quad 1\,)^T, \quad \mathbf{a}_3 = (\,1 \quad 1 \quad 1\,)^T\,; \\
&(b) \quad \mathbf{a}_1 = (\,-1 \quad 2 \quad 2\,)^T, \quad \mathbf{a}_2 = (\,4 \quad -5 \quad -2\,)^T, \quad \mathbf{a}_3 = (\,5 \quad -2 \quad 1\,)^T\,.
\end{aligned}$$

samt QR-faktorisera matrisen $A = (\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,)$.
11. (a) Konstruera en ortonormal bas (dvs. en bas som är ett ON-system) till det plan V i $\mathbf{R}^3$ som har ekvationen $x - y + 2z = 0$.

(b) Räkna ut projektionsmatrisen (för ortogonal projektion) på V.

(c) Räkna också ut speglingen $S\mathbf{x}$ i V av en godtycklig punkt $\mathbf{x} = (\,x \quad y \quad z\,)^T$ (här är S speglingsmatrisen i V).
12. Visa att för varje m/n-matris A med rangen r gäller att matrisen AA^T är symmetrisk och har rangen r.
13. Vilken är volymen av den parallellepiped, som har sina hörn i punkterna $(\,0 \quad 0 \quad 0\,)$, $(\,-1 \quad 2 \quad 2\,)$, $(\,2 \quad -1 \quad 2\,)$ och $(\,2 \quad 2 \quad -1\,)$?
14. (a) Låt A vara en m/n-matris. Visa att A^TA bara har icke-negativa egenvärden.

(b) Låt $\lambda_1 \geq \cdots \geq \lambda_k > 0$ vara en uppräkning av A^TA:s positiva egenvärden och sätt $\sigma_j = +\sqrt{\lambda_j}$ för $j = 1, \ldots, k$ (dessa tal kallas A:s *singulärvärden*). Låt vidare $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ vara ett ON-system av motsvarande egenvektorer. Visa att $\mathbf{v}_j^T A^T A \mathbf{v}_i = {\sigma_j}^2 \delta_{ji}$ för $j = 1, \ldots, k$ och $i = 1, \ldots, n$.

(c) Sätt $\mathbf{u}_j = \sigma_j^{-1} A \mathbf{v}_j$ för $j = 1, \ldots, k$. Visa att $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ är ett ON-system.

(d) Komplettera $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ till en ON-bas $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ i $\mathbf{R}^m$.

(e) Låt U och V vara de ortogonala matriser som innehåller $\mathbf{u}_j$:na resp. $\mathbf{v}_i$:na som kolonner och låt Σ vara den m/n-matris som diagonalt från övre vänstra hörnet har talen $\sigma_1, \ldots, \sigma_k, 0, \ldots$ medan övriga matriselement är nollor. Visa att $A = U\Sigma V^T$. Detta är *singulärvärdesfaktoriseringen* (singulärvärdesdekompositionen) av A.

9. Diagonalisering

Låt A vara en given n/n-matris. Vi ställer oss som uppgift att finna en inverterbar matris B, sådan att

$$B^{-1}AB = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \tag{1}$$

är en diagonalmatris.

Definition 9.1. En matris B, som den i (1), sägs *diagonalisera* A. Matrisen A sägs vara *diagonaliserbar* om det existerar en matris B som diagonaliserar A.

För att få ett grepp om hur matrisen B skall konstrueras, antar vi att vi redan har hittat en matris B sådan att (1) gäller och undersöker hur B då måste se ut:

Genom att multiplicera med B från vänster skriver vi om (1) i formen $AB = BD$. Sedan sätter vi $B = (\,\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n\,)$ och får

$$\begin{aligned} AB = BD \quad &\Longleftrightarrow \quad A(\,\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n\,) = (\,\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n\,)D \\ &\Longleftrightarrow \quad (\,A\mathbf{b}_1 \quad \ldots \quad A\mathbf{b}_n\,) = (\,\lambda_1\mathbf{b}_1 \quad \ldots \quad \lambda_n\mathbf{b}_n\,) \\ &\Longleftrightarrow \quad A\mathbf{b}_i = \lambda_i\mathbf{b}_i \quad (i = 1, \ldots, n)\,. \end{aligned}$$

Av detta ser vi att diagonalelementen λ_i i D kommer att vara egenvärden till A och att kolonnerna $\mathbf{b}_i$ i B är motsvarande egenvektorer till A. Eftersom matrisen B dessutom måste vara inverterbar (dvs. icke-singulär), så måste egenvektorerna $\mathbf{b}_i$ väljas så att att de är linjärt oberoende.

Vi sammanfattar:

Sats 9.1. *En n/n-matris A är diagonaliserbar om och endast om det finns n stycken linjärt oberoende egenvektorer $\mathbf{b}_1, \ldots, \mathbf{b}_n$ till A. Den matris $B = (\,\mathbf{b}_1 \quad \ldots \quad \mathbf{b}_n\,)$, som har dessa egenvektorer som kolonner, är en matris som diagonaliserar A:*

$$B^{-1}AB = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Egenvärdena till A uppträder i D:s diagonal i en ordning som svarar mot egenvektorernas ordning i B.

Anmärkning. Om A:s alla egenvärden $\lambda_1, \ldots, \lambda_n$ är **olika** så finns det n stycken linjärt oberoende egenvektorer till A. Matrisen A är alltså diagonaliserbar i detta fall. Om A är **symmetrisk** så finns det också alltid n linjärt oberoende egenvektorer (också då vissa egenvärden är lika). Alla symmetriska matriser är således diagonaliserbara.

Exempel 9.1. Matrisen

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

har den karakteristiska ekvationen $(2-\lambda)^2 - 4 = 0$, vilket ger egenvärdena 0 och 4. Efter en enkel kalkyl finner vi baser $\{\mathbf{b}_1\}$ respektive $\{\mathbf{b}_2\}$ i $V(0)$ respektive $V(4)$, där $\mathbf{b}_1 = (\,-1 \quad 2\,)^T$ och $\mathbf{b}_2 = (\,1 \quad 2\,)^T$. På grund av Sats 9.1 vet vi att matrisen $B = (\,\mathbf{b}_1 \quad \mathbf{b}_2\,)$ är en matris som diagonaliserar A, så att

$$B^{-1}AB = D = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}.$$

Observera att vi inte behöver utföra matrismultiplikationerna i vänstra ledet för att få fram D utan vi skriver bara ut A:s egenvärden i D:s diagonal i samma ordning som vi har placerat egenvektorerna i B. För kontrollens skull kan man emellertid utföra multiplikationerna i $AB = BD$ för att upptäcka om man gjort räknefel.

En annan matris som diagonaliserar A är t.ex. $B_1 = (\,2\mathbf{b}_1 \quad 3\mathbf{b}_2\,)$, eftersom $2\mathbf{b}_1$ och $3\mathbf{b}_2$ är två linjärt oberoende egenvektorer till A.

Det finns matriser som inte kan diagonaliseras. Vi ger ett exempel på en sådan:

Exempel 9.2. Matrisen

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

har det enda egenvärdet 0 och vektorn $\mathbf{b} = (\,1 \quad 0\,)^T$ bildar en bas i egenrummet $V(0)$. Alla egenvektorer till A är alltså multipler av $\mathbf{b}$. Därför finns det bara **en** linjärt oberoende egenvektor till A. Men vi skulle ju behöva **två** för att kunna skriva ut en inverterbar matris B som diagonaliserar A. Således är matrisen A inte diagonaliserbar.

Det som vi hittills har sagt om diagonalisering gäller för alla matriser men om matrisen är symmetrisk kan vi till och med visa att diagonaliseringen alltid kan göras med hjälp av en ortogonal matris:

Sats 9.2. *Om A är en symmetrisk n/n-matris så existerar det en ortogonal matris Q* (varvid $Q^{-1} = Q^T$), *sådan att*

$$Q^TAQ = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

där $\lambda_1, \ldots, \lambda_n$ är A:s egenvärden.

Bevis. I varje egenrum räknar man ut en bas, som sedan ortonormeras med hjälp av Gram–Schmidt-proceduren till ett ON-system (se exempel 8.8). Unionen av dessa baser i egenrummen bildar sedan en ON-bas $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ i $\mathbf{R}^n$, eftersom egenrummen i detta fall färdigt är parvis ortogonala. Sätt $Q = (\,\mathbf{q}_1 \quad \ldots \quad \mathbf{q}_n\,)$. Då är Q en ortogonal matris, som enligt Sats 9.1 diagonaliserar A. $\diamondsuit$

Exempel 9.3. Den symmetriska matrisen

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

har egenvärdena -2 och 1. Genom att lösa ekvationen $(A + 2I)\mathbf{x} = \mathbf{0}$ hittar vi basvektorerna

$$\mathbf{a}_1 = (\,-1 \quad 1 \quad 0\,)^T \quad \text{och} \quad \mathbf{a}_2 = (\,-1 \quad 0 \quad 1\,)^T$$

i egenrummet $V(-2)$. Dessa ortonormerar vi med hjälp av Gram–Schmidt-proceduren till två ortogonala enhetsvektorer

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}}(\,-1 \quad 1 \quad 0\,)^T, \quad \mathbf{q}_2 = \frac{1}{\sqrt{6}}(\,-1 \quad -1 \quad 2\,)^T.$$

I egenrummet $V(1)$ hittar vi på liknande sätt en enda basvektor $\mathbf{a}_3 = (\,1 \quad 1 \quad 1\,)^T$, som vi normerar till en enhetsvektor

$$\mathbf{q}_3 = \frac{1}{\sqrt{3}}(\,1 \quad 1 \quad 1\,)^T$$

(steg 2 i Gram–Schmidt-proceduren!). Om vi nu sätter

$$Q = (\,\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3\,) = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix},$$

så är Q en ortogonal matris som diagonaliserar A:

$$Q^TAQ = D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

På grund av att m:te potenser av en diagonalmatris D fås genom att diagonalelementen i D upphöjs till potensen m, kan en diagonalisering $B^{-1}AB = D$ av A användas till att på ett enkelt sätt räkna ut potenser (och funktioner) av A:

$$\begin{aligned} A &= BDB^{-1} \\ A^2 &= BDB^{-1}BDB^{-1} = BD^2B^{-1} \\ &- - - - \\ A^m &= BD^mB^{-1}. \end{aligned}$$

Ur detta följer att om $p(t) = a_0 + a_1t + \cdots + a_rt^r$ är ett polynom och om vi sätter $p(A) = a_0I + a_1A + \cdots + a_rA^r$, så är

$$p(A) = Bp(D)B^{-1} = B \begin{pmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_n) \end{pmatrix} B^{-1}.$$

Exempel 9.4. För t.ex. den matris som vi diagonaliserade i exempel 1 i detta avsnitt, är

$$p(A) = \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} p(0) & 0 \\ 0 & p(4) \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}^{-1}$$
$$= \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} p(0) & 0 \\ 0 & p(4) \end{pmatrix} \frac{1}{4} \begin{pmatrix} -2 & 1 \\ 2 & 1 \end{pmatrix}.$$

för varje polynom p.

Övningsuppgifter

1. Avgör om följande matriser är diagonaliserbara:

$$(a) \quad \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}, \qquad (b) \quad \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \qquad (c) \quad \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

2. Räkna ut A^n genom att först diagonalisera matrisen

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

Bestäm också en kvadratrot ur A, dvs. en matris X sådan att $X^2 = A$.

3. Diagonalisera den symmetriska matrisen

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

med hjälp av en ortogonal matris.

4. Diagonalisera den symmetriska matrisen

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \qquad (b \neq 0),$$

med hjälp av en ortogonal matris samt räkna ut matrisen A^n.

5. Bestäm en 2/2-matris A med 2 och -3 som egenvärden och $(-1 \quad 2)^T$ respektive $(1 \quad 1)^T$ som motsvarande egenvektorer.

6. I en djurpopulation lever en individ högst n år. Vi låter $x_i^{(k)}$ beteckna antalet honor med åldern i år k $(k \geq 0)$. Vidare låter vi f_i beteckna den bråkdel av honorna med åldern i som överlever till nästa år och b_i må beteckna medelantalet ungar av honkön som en hona med åldern i föder ett visst år. Med $\mathbf{x}^{(k)} = \left(x_0^{(k)} \ldots x_n^{(k)}\right)^T$ betecknar vi *åldersfördelningsvektorn* år k. (a) Skriv ut den matris A (*Leslie-matrisen*) som ger följande års åldersfördelningsvektor enligt formeln $\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)}$. (b) Visa att $\mathbf{x}^{(k)}$ **kan** vara konstant (år efter år) om och endast om A har talet 1 som egenvärde. (c) Antag att en viss skalbagge har Leslie-matrisen

$$A = \begin{pmatrix} 0 & 4 & 10 \\ 0,2 & 0 & 0 \\ 0 & 0,1 & 0 \end{pmatrix}.$$

Är en konstant åldersfördelning möjlig? Har arten stora chanser att överleva?

7. (Exempel på en *Markovkedja*) Mor Stava har egenheten att vara glad eller ledsen en hel dag i sträck. Om hon är glad en viss dag, är sannolikheten för att hon skall vara glad respektive ledsen följande dag $0,6$ och $0,4$. Hon orkar inte vara ledsen särskilt länge, så att om hon är ledsen en viss dag så är sannolikheten för att hon är glad resp. ledsen följande dag $0,8$ och $0,2$. Om vektorn $(\,p \quad q\,)^T$ anger sannolikheterna p och $q = 1 - p$ för att Stava är glad resp. ledsen dag 0, så anger (enligt elementär sannolikhetslära) vektorn

$$\begin{pmatrix} 0,6p + 0,8q \\ 0,4p + 0,2q \end{pmatrix} = \begin{pmatrix} 0,6 & 0,8 \\ 0,4 & 0,2 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = A \begin{pmatrix} p \\ q \end{pmatrix}$$

motsvarande sannolikheter för dag 1 samt $A^n \begin{pmatrix} p \\ q \end{pmatrix}$ motsvarande sannolikheter för dag n. En hur stor andel av sina levnadsdagar är mor Stava glad? (Ledning: Diagonalisera A och låt $n \to \infty$.)

10. Linjära operatorer

En operator är en funktion f, som avbildar en vektor $\mathbf{x}$ på en annan vektor $f(\mathbf{x})$. Funktionsvärdet $f(\mathbf{x})$ och argumentet $\mathbf{x}$ kan tillhöra olika vektorrum. Vi kommer bara att befatta oss med operatorer som är linjära. Linjära operatorer kan nämligen associeras direkt med matriser.

Följande exempel på en linjär operator kan kallas typexemplet för alla linjära operatorer:

Exempel 10.1. Låt A vara en m/n-matris och låt en funktion (dvs. en operator) $T : \mathbf{R}^n \to \mathbf{R}^m$ vara definierad genom

$$T(\mathbf{x}) = A\mathbf{x} \qquad \text{för varje } \mathbf{x} \in \mathbf{R}^n\,.$$

Räknereglerna för matrismultiplikation ger att för denna funktion gäller

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{och} \\ T(\lambda\mathbf{x}) &= A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda T(\mathbf{x})\,. \end{aligned}$$

En operator som har dessa egenskaper sägs vara linjär:

Definition 10.1. Låt E och F vara vektorrum. En *linjär operator* T från E till F är en funktion $T : E \to F$, som är definierad för varje argument $\mathbf{x} \in E$, antar sina värden i F och dessutom uppfyller linearitetsvillkoren

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T(\mathbf{x}) + T(\mathbf{y}) && \text{för varje } \mathbf{x},\, \mathbf{y} \in E, \\ T(\lambda\mathbf{x}) &= \lambda T(\mathbf{x}) && \text{för varje } \mathbf{x} \in E,\ \lambda \in \mathbf{R}. \end{aligned}$$

Anmärkning. Dessa linearitetsvillkor är uppfyllda också för sammansättningar: Om S och T är linjära operatorer från E till F och om också operatorn $U : F \to G$ är linjär, så är de sammansatta operatorerna $S + T : E \to F$ och $U \circ T : E \to G$ linjära (övningsuppgift 1). Dessa sammansättningar är definierade genom

$$\begin{aligned} (S+T)(\mathbf{x}) &= S(\mathbf{x}) + T(\mathbf{x})\,, \\ (U \circ T)(\mathbf{x}) &= U(T(\mathbf{x}))\,. \end{aligned}$$

Exempel 10.2. Låt P beteckna vektorrummet av alla polynom. Då är deriveringsoperatorn $D : P \to P$, vilkens värde för varje polynom $p \in P$ är definierad som derivatan

$$(D(p))(x) = p'(x)\,,$$

en linjär operator. Välkända deriveringsregler säger ju att

$$\begin{aligned} D(p_1 + p_2) &= p_1' + p_2' = D(p_1) + D(p_2)\,, \\ D(\lambda p) &= \lambda p' = \lambda D(p)\,. \end{aligned}$$

Exempel 10.3. Ett annat exempel på en linjär operator på samma vektorrum P, är den operator $T : P \to P$, som är definierad genom

$$T(p)\,(x) = xp(x)\,.$$

Eftersom $x(p_1(x) + p_2(x)) = xp_1(x) + xp_2(x)$ och $x(\lambda p(x)) = \lambda(xp(x))$, är linearitetsvillkoren uppfyllda.

Sats 10.1. *För en linjär operator $T : E \to F$ från ett vektorrum E till ett vektorrum F gäller:*

(*i*) $T(\mathbf{0}) = \mathbf{0}\,.$

(*ii*) *Om vektorerna $\mathbf{x}_1, \ldots, \mathbf{x}_p$ är linjärt beroende i E så är $T(\mathbf{x}_1), \ldots, T(\mathbf{x}_p)$ linjärt beroende i F.*

Bevis. (*i*) För ett godtyckligt $\mathbf{x} \in E$ är $T(\mathbf{0}) = T(0\mathbf{x}) = 0T(\mathbf{x}) = \mathbf{0}$.
(*ii*) Om $\mathbf{x}_1, \ldots, \mathbf{x}_p$ är linjärt beroende så existerar icke-triviala koefficienter c_i sådana att $c_1\mathbf{x}_1 + \cdots + c_p\mathbf{x}_p = \mathbf{0}$. Men ur linearitetsvillkoren följer då att

$$c_1T(\mathbf{x}_1) + \cdots c_pT(\mathbf{x}_p) = T(c_1\mathbf{x}_1 + \cdots + c_p\mathbf{x}_p) = T(\mathbf{0}) = \mathbf{0}\,.$$

för samma koefficienter. Alltså är $T(\mathbf{x}_1), \ldots, T(\mathbf{x}_p)$ linjärt beroende. $\diamondsuit$

Vi betraktar ytterligare några exempel på linjära operatorer:

Exempel 10.4. Operatorn $T : \mathbf{R}^3 \to \mathbf{R}^3$,

$$T\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T = \begin{pmatrix} -x_1 & -x_2 & x_3 \end{pmatrix}^T,$$

speglar vektorn $\mathbf{x} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T$ i x_3-axeln.

Exempel 10.5. Operatorn $T : \mathbf{R}^3 \to \mathbf{R}^3$,

$$T\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T = \begin{pmatrix} x_1 & x_2 & -x_3 \end{pmatrix}^T,$$

speglar vektorn $\mathbf{x} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T$ i x_1x_2-planet.

Anmärkning. Om $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ är en bas i ett vektorrum E så blir en linjär operator $T : E \to F$ entydigt definierad om vi fastslår värdena $T(\mathbf{a}_1), \ldots, T(\mathbf{a}_n)$. För ett godtyckligt $\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$ är nämligen då med nödvändighet $T(\mathbf{x}) = x_1T(\mathbf{a}_1) + \ldots + x_nT(\mathbf{a}_n)$ (lineariteten hos T är lätt att verifiera p.g.a. att detta uttryck för $T(\mathbf{x})$ är linjärt i koordinaterna x_i).

Exempel 10.6. Den linjära operator $T : \mathbf{R}^2 \to \mathbf{R}^2$, som vrider varje vektor en vinkel θ moturs kring origo, kan enligt anmärkningen ovan definieras genom

$$T(\mathbf{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \quad \text{och} \quad T(\mathbf{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

I exemplen 10.4, 10.5 och 10.6 finns det en matris A sådan att $T(\mathbf{x}) = A\mathbf{x}$. Låt oss t.ex i exempel 10.4 sätta $\mathbf{y} = T(\mathbf{x})$, där $\mathbf{y} = (\,y_1 \quad y_2 \quad y_3\,)^T$. Då är

$$\begin{cases} y_1 = -x_1 \\ y_2 = -x_2\,, \\ y_3 = x_3 \end{cases} \quad \text{dvs.} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A\mathbf{x}\,.$$

För operatorn T i exempel 10.6 får ur $\mathbf{y} = T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2)$ att

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Detta är ingen tillfällighet. Vi kommer att se att mot varje linjär operator T från ett n-dimensionellt vektorrum till ett m-dimensionellt vektorrum entydigt svarar en m/n-matris samt att T entydigt kan rekonstrueras ur denna matris. Denna motsvarighet mellan operator och matris är dock beroende av valet av baser i vektorrummen. Vi skall reda ut detta närmare:

Sambandet mellan linjära operatorer och matriser

Låt E och F vara vektorrum och antag att $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ respektive $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ är (numrerade) baser i dessa. Låt dessutom $T : E \to F$ vara en linjär operator.

Varje $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$ är entydigt definierat av sina koordinater, som vi sammanfattar i en *koordinatvektor*

$$X = (\,x_1 \quad \ldots \quad x_n\,)^T\,.$$

På grund av lineariteten är

(1) $$T(\mathbf{x}) = x_1 T(\mathbf{b}_1) + \cdots + x_n T(\mathbf{b}_n) = \sum_{k=1}^{n} x_k T(\mathbf{b}_k)\,.$$

Varje $T(\mathbf{b}_k)$ i detta uttryck är en linjärkombination av basvektorerna i F,

(2) $$T(\mathbf{b}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{v}_i\,, \quad (k = 1, \ldots, n)\,.$$

Definition 10.1. Den matris som koefficienterna i (2) bildar,

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

kallas *matrisen för T i baserna* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ och $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$.

Anmärkning. En matris A definierar entydigt en operator T då (de numrerade) baserna är fastslagna. Genom (2) är ju $T(\mathbf{b}_k)$ definierat för $k = 1, \ldots, n$ och genom (1) är sedan $T(\mathbf{x})$ definierat för varje $\mathbf{x}$.

Exempel 10.7. Låt operatorn $T : \mathbf{R}^2 \to \mathbf{R}^2$ vara definierad genom

$$T(\mathbf{x}) = A\mathbf{x}, \qquad \text{där} \qquad A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Då är matrisen för T i den naturliga basen $\{\mathbf{e}_1, \mathbf{e}_2\}$ just matrisen A, ty

$$\begin{aligned} T(\mathbf{e}_1) &= (\,1 \quad 2\,)^T = \mathbf{e}_1 + 2\mathbf{e}_2 \\ T(\mathbf{e}_2) &= (\,-1 \quad 1\,)^T = -\mathbf{e}_1 + \mathbf{e}_2\,. \end{aligned}$$

Betrakta nu den bas i $\mathbf{R}^2$, som består av vektorerna $\mathbf{v}_1 = (\,1 \quad 0\,)^T$ och $\mathbf{v}_2 = (\,1 \quad 1\,)^T$. För att få fram matrisen för T i denna nya bas, räknar vi ut

$$\begin{aligned} T(\mathbf{v}_1) &= \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -\mathbf{v}_1 + 2\mathbf{v}_2 \\ T(\mathbf{v}_2) &= \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = -3\mathbf{v}_1 + 3\mathbf{v}_2 \end{aligned}$$

och avläser att matrisen för T i basen $\{\mathbf{v}_1, \mathbf{v}_2\}$ är

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

Exempel 10.8. Låt P_k betecknar vektorrummet av polynom av högst graden k ($k = 2, 3$). Betrakta i P_k den bas som består av polynomen $p_j(x) = x^j$ ($j = 0, \ldots, k$). För operatorn $T : P_2 \to P_3$, $T(p)\,(x) = xp(x)$, gäller

$$\begin{aligned} T(p_0)\,(x) &= x\,p_0(x) = x \;\; = p_1(x) \\ T(p_1)\,(x) &= x\,p_1(x) = x^2 = p_2(x) \\ T(p_2)\,(x) &= x\,p_2(x) = x^3 = p_3(x)\,. \end{aligned}$$

Därför är

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

matrisen för T i baserna $\{p_0, p_1, p_2\}$ och $\{p_0, p_1, p_2, p_3\}$.

Vi går nu tillbaka till det allmänna fallet. Låt $T : E \to F$ vara den linjära operator, som vi betraktade i början av detta avsnitt och som har matrisen A i definition 10.1. Avsikten är att skriva likheten $\mathbf{y} = T(\mathbf{x})$ i matrisform. Därför inför vi beteckningarna

$$X = (\,x_1 \quad \ldots \quad x_n\,)^T \qquad \text{och} \qquad Y = (\,y_1 \quad \ldots \quad y_m\,)^T$$

för koordinatvektorerna för $\mathbf{x}$ och $\mathbf{y}$ i respektive bas. Genom insättning av i tur och ordning (1) och (2) fås

$$\mathbf{y} = T(\mathbf{x}) = \sum_{k=1}^{n} x_k T(\mathbf{b}_k) = \sum_{k=1}^{n} x_k \sum_{i=1}^{m} a_{ik}\mathbf{v}_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} x_k \right) \mathbf{v}_i\,.$$

Koefficienterna för basvektorerna $\mathbf{v}_i$ måste här vara koordinaterna för $\mathbf{y}$,

$$y_i = \sum_{k=1}^{n} a_{ik} x_k, \qquad \text{dvs.} \qquad Y = AX\,.$$

Omvänt kan likheten $\mathbf{y} = T(\mathbf{x})$ entydigt härledas ur $Y = AX$, då ju koordinatvektorerna entydigt bestämmer $\mathbf{x}$ och $\mathbf{y}$ och operatorn T, som vi har sett, entydigt kan rekonstrueras ur A.

Vi sammanfattar:

Sats 10.2. *Antag att $T : E \to F$ är en linjär operator mellan vektorrum E och F med baser $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ respektive $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. Då gäller ekvivalensen*

$$\mathbf{y} = T(\mathbf{x}) \quad \Longleftrightarrow \quad Y = AX\,,$$

där X och Y betecknar koordinatvektorerna för $\mathbf{x}$ och $\mathbf{y}$ i respektive bas och A är matrisen för T i de nämnda baserna.

Anmärkning. Observera att den *identiska operatorn* $Id : E \to E$, definierad genom $Id(\mathbf{x}) = \mathbf{x}$, har matrisen I i varje bas i E.

Matrisen för sammansättningar

Låt $T : E \to F$ och $S : F \to G$ vara linjära operatorer mellan ändligtdimensionella vektorrum och betrakta sammansättningen $S \circ T : E \to G$. Vi antar att baser har valts i E, F och G samt att T, S och $S \circ T$ har matriserna B, A respektive C i dessa baser. Enligt Sats 10.2 gäller ekvivalenserna

$$\begin{aligned} \mathbf{y} = T(\mathbf{x}) \quad &\Longleftrightarrow \quad Y = BX\,, \\ \mathbf{z} = S(\mathbf{y}) \quad &\Longleftrightarrow \quad Z = AY\,, \\ \mathbf{z} = (T \circ S)(\mathbf{x}) \quad &\Longleftrightarrow \quad Z = CX\,, \end{aligned}$$

där X, Y och Z är koordinatvektorerna för $\mathbf{x}$, $\mathbf{y}$ och $\mathbf{z}$ i respektive bas. Insättning av $Y = BX$ i $Z = AY$ ger $Z = ABX$. Således är

$$CX = ABX \qquad \text{för varje } X,$$

dvs. $C = AB$. Vi har alltså visat:

Sats 10.3. *Antag att baser har valts i de ändligtdimensionella vektorrummen E, F och G. Om de linjära operatorerna $T : E \to F$ och $S : F \to G$ har matriserna B och A i respektive baser så har $S \circ T$ matrisen AB i de valda baserna i E och G.*

Bastransformationer

Som vi har sett, är matrisen för en operator $T : E \to F$ beroende av vilka baser man använder i E och F. Då man byter baser, dvs. utför en bastransformation, ändras matrisen på ett visst sätt. Vi skall studera detta närmare för det fall att operatorn T har både argument och värde i samma vektorrum. Det allmänna fallet behandlas på liknande sätt men är lite mera komplicerat och inte lika viktigt (se sid. 135).

Låt alltså $T : E \to E$ vara en linjär operator med matrisen $A = (a_{ik})$ i basen $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ och matrisen $A' = (a'_{ik})$ i den nya basen $\{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$:

$$T(\mathbf{b}_k) = \sum_{i=1}^{n} a_{ik}\mathbf{b}_i\,, \qquad T(\mathbf{b}'_k) = \sum_{i=1}^{n} a'_{ik}\mathbf{b}'_i\,.$$

Vi söker en formel som anger sambandet mellan A och A'. Det visar sig att vi lätt kan få fram en sådan formel, om vi först utreder hur koordinatvektorer förändras vid basbyten.

Vi skriver de nya basvektorerna som linjärkombinationer av de gamla,

$$\begin{aligned} \mathbf{b}'_1 &= c_{11}\mathbf{b}_1 + \cdots + c_{n1}\mathbf{b}_n \\ &- \ - \ - \ - \ - \ - \\ \mathbf{b}'_n &= c_{1n}\mathbf{b}_1 + \cdots + c_{nn}\mathbf{b}_n\,, \end{aligned}$$

eller i kortare form,

$$\mathbf{b}'_k = \sum_{i=1}^{n} c_{ik}\mathbf{b}_i \qquad (k = 1, \ldots, n)\,. \tag{3}$$

Vi kan då säga att basbytet beskrivs av *basbytesmatrisen* (eller *bastransformationsmatrisen*)

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}.$$

Låt $\mathbf{x}$ vara en godtycklig vektor i E samt låt $X = (x_i)$ och $X' = (x'_i)$ vara koordinatvektorena för detta $\mathbf{x}$ i den "gamla" respektive den nya basen. Då ger (3) att

$$\mathbf{x} = \sum_{k=1}^{n} x'_k\mathbf{b}'_k = \sum_{k=1}^{n} x'_n \sum_{i=1}^{n} c_{ik}\mathbf{b}_i = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} c_{ik}x'_k\right)\mathbf{b}_i\,,$$

ur vilket sambandet $x_i = \sum_{k=1}^{n} c_{ik}x'_k$, dvs. $X = CX'$, kan avläsas.

Sats 10.4. *Låt E vara ett vektorrum med en ursprunglig bas $\{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ och en ny bas $\{\mathbf{b}'_1, \cdots, \mathbf{b}'_n\}$, relaterade genom (3). Då är koordinatvektorerna X respektive X' för en given vektor $\mathbf{x}$ i E i den ursprungliga och den nya basen relaterade genom*

$$X = CX'\,,$$

där C är basbytesmatrisen, som består av koefficienterna i (3).

Anmärkning. $R(C) = \mathbf{R}^n$, eftersom varje $X \in \mathbf{R}^n$ kan fås som en produkt CX'. Följaktligen är C icke-singulär och därmed inverterbar. Alltså är $X' = C^{-1}X$.

Exempel 10.9. Antag att en vektor $\mathbf{x}$ i $\mathbf{R}^2$ har koordinatvektorn $X = (1 \quad 3)^T$ i viss (för oss okänd) bas $\{\mathbf{b}_1, \mathbf{b}_2\}$. Vi inför en ny (okänd) bas genom att sätta

$$\begin{aligned} \mathbf{b}_1' &= 2\mathbf{b}_1 - 3\mathbf{b}_2 \,, \\ \mathbf{b}_2' &= \ \ \mathbf{b}_1 + \ \mathbf{b}_2 \,. \end{aligned}$$

Då är basbytesmatrisen

$$C = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}, \qquad \text{som har inversen} \qquad C^{-1} = \frac{1}{5}\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}.$$

Koordinatvektorn för $\mathbf{x}$ i den nya basen blir alltså

$$X' = C^{-1}X = \frac{1}{5}\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -2 \\ 9 \end{pmatrix}.$$

Vi går nu tillbaka till problemet att finna ett samband mellan matriserna A och A' för operatorn T i de två baserna. Låt vektorerna $\mathbf{x}$ och $\mathbf{y}$ i likheten $\mathbf{y} = T(\mathbf{x})$ ha koordinatvektorer X och Y respektive X' och Y' i den ursprungliga respektive den nya basen. Enligt Sats 10.4 är nu $X = CX'$ och $Y = CY'$ och vi får de ekvivalenta relationerna

$$\mathbf{y} = T(\mathbf{x}) \quad \Leftrightarrow \quad Y = AX \quad \Leftrightarrow \quad CY' = ACX' \quad \Leftrightarrow \quad Y' = (C^{-1}AC)X' \,,$$

ur vilka vi kan avläsa den sökta formeln $A' = C^{-1}AC$.

Sats 10.5. *Låt E vara ett vektorrum med en ursprunglig bas $\{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ och en ny bas $\{\mathbf{b}_1', \cdots, \mathbf{b}_n'\}$ och låt baserna vara relaterade genom (3). Om $T : E \to E$ är en linjär operator med matriserna A och A' i den ursprungliga respektive den nya basen, så är dessa matriser relaterade genom*

$$A' = C^{-1}AC \,. \tag{4}$$

Exempel 10.10. Betrakta operatorn $T : \mathbf{R}^3 \to \mathbf{R}^3$, definierad genom

$$T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Då är A matrisen för T i den naturliga basen $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Vi skall emellertid bestämma matrisen för T i den bas som består av vektorerna

$$\mathbf{b}_1 = (1 \quad 2 \quad 0)^T \,, \quad \mathbf{b}_2 = (0 \quad 1 \quad 2)^T \,, \quad \mathbf{b}_3 = (2 \quad 0 \quad 1)^T \,.$$

Då är

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{e}_1 + 2\mathbf{e}_2 \\ \mathbf{b}_2 &= \mathbf{e}_2 + 2\mathbf{e}_3 \\ \mathbf{b}_3 &= 2\mathbf{e}_1 + \mathbf{e}_3 \,. \end{aligned}$$

Ur detta avläser vi basbytesmatrisen

$$C = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \text{ och får genom invertering } C^{-1} = \frac{1}{9}\begin{pmatrix} 1 & 4 & 2 \\ -2 & 1 & 4 \\ 4 & -2 & 1 \end{pmatrix}.$$

Matrisen $A' = C^{-1}AC$ för T i basen $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ blir nu

$$A' = \frac{1}{9}\begin{pmatrix} 1 & 4 & 2 \\ -2 & 1 & 4 \\ 4 & -2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 18 & 28 & 23 \\ 7 & 13 & -2 \\ 4 & 1 & 13 \end{pmatrix}.$$

I det allmänna fallet då man har en operator $T : E \to F$ och olika baser $\mathbf{b}_1, \dots, \mathbf{b}_n\}$ och $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ i E respektive F så kan man göra olika bastransformationer i de två vektorrummen:

Låt $\{\mathbf{b}'_1, \dots, \mathbf{b}'_n\}$ och $\{\mathbf{v}'_1, \dots, \mathbf{v}'_m\}$ vara nya baser i E respektive F. Betrakta likheten $\mathbf{y} = T(\mathbf{x})$ och låt X, Y respektive X' och Y' vara koordinatvektorerna för $\mathbf{x}$ och $\mathbf{y}$ i de ursprungliga respektive de nya baserna. Låt C och D vara basbytesmatriserna i E respektive F och låt A vara matrisen för T i de ursprungliga baserna. Då är enligt Sats 10.4

$$X = CX' \quad \text{och} \quad Y = DY' \,.$$

Genom insättning av dessa i $Y = AX$ fås ekvationen $DY' = ACX'$, vilket ger att $Y' = (D^{-1}AC)X'$. Matrisen för T i de nya baserna blir alltså

$$A' = D^{-1}AC \,.$$

Anmärkning. Eftersom den operation (4), som görs vid en bastransformation, är identisk med den som görs vid diagonalisering, så kan vi tolka diagonalisering av en matris A på följande sätt: Om A är diagonaliserbar, så kan matrisen för den linjära operatorn $T(\mathbf{x}) = A\mathbf{x}$ fås diagonal genom att man väljer ett lämpligt nytt koordinatsystem, dvs. väljer en ny bas. Vi har sett att en diagonalisering alltid är möjlig om A är symmetrisk och att det nya koordinatsystemet då alltid kan väljas ortogonalt. Transformationen till den nya koordinatvektorn, $X' = Q^{-1}X$, förmedlas i detta fall av en ortogonal matris Q. Om A är en 2/2- eller 3/3-matris, kan man t.o.m. visa att Q alltid kan väljas att vara en rotationsmatris.

Inversa operatorer

Antag att $T : E \to E$ är en linjär operator. Eftersom T är en vanlig funktion, existerar den *inversa operatorn* T^{-1} (den inversa funktionen) om och endast om T är bijektiv, dvs. både surjektiv och injektiv. Vi repeterar dessa begrepp:

(i) T sägs vara surjektiv om $\{T(\mathbf{x}) \mid \mathbf{x} \in E\} = E$;

(ii) T sägs vara injektiv om $T(\mathbf{x}) = T(\mathbf{y}) \quad \Rightarrow \quad \mathbf{x} = \mathbf{y}$.

Om operatorn T^{-1} existerar, är den automatiskt linjär (se övningsuppgift 2).

Vi skall strax visa att villkoren (i) och (ii) i själva verket är ekvivalenta för en linjär operator om E är ändligtdimensionellt. Antag nu att $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ är en (ändlig) bas i E och låt A och B vara matriserna för T respektive T^{-1} samt X och Y koordinatvektorerna för $\mathbf{x}$ respektive $\mathbf{y}$ i denna bas. Notera först att relationen $T \circ T^{-1} = Id$ enligt Sats 10.3 medför att $AB = I$, varför $B = A^{-1}$. Följande kedja av ekvivalenser gäller:

$$
\begin{array}{rcl}
T \text{ är surjektiv} & \Leftrightarrow & \text{alla } \mathbf{y} \in E \text{ är av formen } \mathbf{y} = T(\mathbf{x}) \\
& \Leftrightarrow & \text{alla } Y \in \mathbf{R}^n \text{ är av formen } Y = AX \\
& \Leftrightarrow & R(A) = \mathbf{R}^n \\
& \Leftrightarrow & \dim R(A) = n \\
A^{-1} \text{ existerar} & \Leftrightarrow & \dim N(A) = 0 \\
& \Leftrightarrow & AX = 0 \text{ bara om } X = 0 \\
& \Leftrightarrow & T(\mathbf{x}) = \mathbf{0} \text{ bara om } \mathbf{x} = \mathbf{0} \\
& \Leftrightarrow & T \text{ är injektiv}
\end{array}
$$

Den sista ekvivalensen bör kanske motiveras: Om vi vet att $T(\mathbf{x}) = \mathbf{0}$ bara om $\mathbf{x} = \mathbf{0}$, så gäller:

$$T(\mathbf{x}) = T(\mathbf{y}) \quad \Rightarrow \quad T(\mathbf{x} - \mathbf{y}) = T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} - \mathbf{y} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{y}\,.$$

Alltså är T injektiv. Omvänt gäller: Om T är injektiv så följer ur $T(\mathbf{x}) = \mathbf{0}$ $(= T(\mathbf{0}))$ att $\mathbf{x} = \mathbf{0}$.

Vi sammanfattar allt detta:

Sats 10.6. *Antag att $T : E \to E$ är en linjär operator och att T har matrisen A i en bas $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Följande påståenden är då ekvivalenta:*

(i) *T är surjektiv;*

(ii) *T är injektiv;*

(iii) *T^{-1} existerar;*

(iv) *A^{-1} existerar. Härvid är A^{-1} matrisen för T^{-1} i basen $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.*

Exempel 10.11. Betrakta den linjära operator $T : P_2 \to P_2$ på vektorrummet P_2 av polynom av högst graden 2, vilken är definierad av uttrycket $T = D^2 + D + I$, där D betecknar deriveringsoperatorn (se exempel 10.2). Om vi i P_2 betraktar den bas som bildas av $p_0(x) = 1$, $p_1(x) = x$ och $p_2(x) = x^2$ och räknar ut

$$\begin{aligned}
T(p_0)(x) &= 1 = p_0(x)\,,\\
T(p_1)(x) &= 1 + x = p_0(x) + p_1(x)\,,\\
T(p_2)(x) &= 2 + 2x + x^2 = 2p_0(x) + 2p_1(x) + p_2(x)\,,
\end{aligned}$$

så kan vi avläsa matrisen för T,

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

Eftersom denna är inverterbar med inversen

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

kan vi med stöd av Sats 10.6 sluta oss till att den inversa operatorn $T^{-1} : P_2 \to P_2$ existerar. Detta innebär att för varje polynom $g \in P_2$ har differentialekvationen $T(p) = p'' + p' + p = g$ exakt en lösning $p \in P_2$. Denna lösning kan räknas ut med matrismetoder: Om t.ex. $g(x) = 1 + x^2$, så har g koordinatvektorn $G = (1 \quad 0 \quad 1)^T$. Koordinatvektorn X för p fås då som lösningen till $AX = G$. Man finner att $X = (1 \quad -2 \quad 1)^T$. Alltså är $p(x) = 1 - 2x + x^2$ den enda lösningen i P_2 till differentialekvationen $p''(x) + p'(x) + p(x) = 1 + x^2$.

Övningsuppgifter

1. Bevisa att om S och T är linjära operatorer, så är också $T \circ S$ en linjär operator.
2. Visa att om $T : E \to E$ är en omvändbar linjär operator (dvs. om T^{-1} existerar) så är också T^{-1} en linjär operator.
3. Kan operatorn T vara linjär om

   (a) $T(1 \quad 0 \quad 1) = (-1 \quad 1 \quad 0)$, $\quad T(0 \quad 2 \quad 1) = (0 \quad 2 \quad 0)$,
   och $\quad T(1 \quad 2 \quad -1) = (1 \quad 0 \quad 1)$?

   (b) $T(2 \quad 4 \quad 2) = (1 \quad 0 \quad 4)$, $\quad T(-1 \quad 0 \quad -2) = (1 \quad 2 \quad 1)$,
   och $\quad T(1 \quad 4 \quad 0) = (2 \quad 2 \quad 2)$?

   (c) $T(1 \quad 1 \quad 1) = (1 \quad 0 \quad -1)$, $\quad T(3 \quad 1 \quad -1) = (0 \quad 1 \quad 1)$,
   och $\quad T(2 \quad 0 \quad -2) = (-1 \quad 1 \quad 2)$?

4. Vektorrummen E och F har baserna $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ respektive $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ och $T : E \to F$ är en linjär operator, sådan att $T(\mathbf{b}_k) = \mathbf{v}_k + \mathbf{v}_{k+1}$, $1 \le k \le 3$. Ange matrisen för T i de nämnda baserna.

5. Låt $T : \mathbf{R}^3 \to \mathbf{R}^3$ vara den linjära operator som i den bas som bildas av $\mathbf{v}_1 = (1 \quad 1 \quad 1)^T$, $\mathbf{v}_2 = (1 \quad 1 \quad 0)^T$ och $\mathbf{v}_3 = (0 \quad 1 \quad -1)^T$ har matrisen

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

Bestäm alla vektorer $\mathbf{x} \in \mathbf{R}^3$, som är sådana att $T(\mathbf{x}) = \mathbf{x}$.

6. Antag att $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ är en bas i $\mathbf{R}^3$. Låt en ny bas $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ vara definierad genom

$$\begin{aligned} \mathbf{b}'_1 &= \mathbf{b}_1 \phantom{+ 2\mathbf{b}_2} + \mathbf{b}_3 \\ \mathbf{b}'_2 &= \phantom{\mathbf{b}_1 +} \mathbf{b}_2 \\ \mathbf{b}'_3 &= \mathbf{b}_1 + 2\mathbf{b}_2 - \mathbf{b}_3 \end{aligned} \quad . \quad \text{Sätt} \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 2 & 0 & 3 \end{pmatrix}.$$

(a) Visa att $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ är en bas i $\mathbf{R}^3$.

(b) Bestäm matrisen för den linjära operatorn $T : \mathbf{R}^3 \to \mathbf{R}^3$ i basen $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ om A är matrisen för T i basen $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

7. Låt $\{\mathbf{b}_1, \mathbf{b}_2\}$ vara en given bas i $\mathbf{R}^2$. Bestäm en ny bas $\{\mathbf{v}_1, \mathbf{v}_2\}$ i $\mathbf{R}^2$ sådan att vektorerna $2\mathbf{b}_1 + 3\mathbf{b}_2$ och $4\mathbf{b}_1 - \mathbf{b}_2$ i denna har koordinaterna 1, 1 respektive 1, -1.

8. Betrakta en linjär operator $T : \mathbf{R}^n \to \mathbf{R}^n$ med matrisen A i basen $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Vi gör ett basbyte till en ny bas $\{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ i $\mathbf{R}^n$ enligt formel (3) där basbytesmatrisen är C. Visa att T har samma matris A i den nya basen om och endast om $AC = CA$.

9. Låt $D : P_3 \to P_2$, $(D(p))(x) = p'(x)$, vara deriveringsoperatorn (P_n är vektorrummet av polynom av högst graden n). Betrakta följande baser i de berörda vektorrummen:

$$B_1 = \{1, x, x^2, x^3\}, \qquad C_1 = \{1, x, x^2\},$$
$$B_2 = \{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}, \qquad C_2 = \{1, 1+x, 1+x+x^2\}.$$

Bestäm matrisen för D i baserna

(a) B_1 och C_1, (b) B_1 och C_2, (c) B_2 och C_1, (d) B_2 och C_2.

10. Mellan baserna $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ och $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ i $\mathbf{R}^3$ råder sambandet

$$\begin{aligned} \mathbf{b}'_1 &= \mathbf{b}_1 + 2\mathbf{b}_2 - 3\mathbf{b}_3 \\ \mathbf{b}'_2 &= \mathbf{b}_1 + \mathbf{b}_2 \\ \mathbf{b}'_3 &= \phantom{\mathbf{b}_1 + 2} \mathbf{b}_2 + \mathbf{b}_3 . \end{aligned}$$

En vektor $\mathbf{x}$ har i basen $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ koordinatvektorn $(x_1 \quad x_2 \quad x_3)^T$ och i basen $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ koordinatvektorn $(x'_1 \quad x'_2 \quad x'_3)^T$. Ange koordinaterna x_1, x_2 och x_3 uttryckta i x'_1, x'_2 och x'_3.

11. Antag att vektorerna $\mathbf{b}_1$, $\mathbf{b}_2$ och $\mathbf{b}_3$ har koordinatvektorerna

$$(1 \quad 0 \quad 0)^T, \quad (2 \quad 1 \quad 0)^T, \quad \text{respektive} \quad (3 \quad 2 \quad 1)^T$$

i en bas $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ i ett vektorrum E. Visa att $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ är en bas i E. Bestäm koordinaterna för vektorn $\mathbf{b}_1 + \mathbf{b}_2$ i basen $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$. Bestäm i basen $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ koordinaterna för den vektor vars koordinatvektor i basen $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ är $(\,2 \quad 0 \quad 7\,)^T$.

12. En linjär operator $T : \mathbf{R}^2 \to \mathbf{R}^3$ har matrisen

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

i de naturliga baserna i $\mathbf{R}^2$ och $\mathbf{R}^3$. För en linjär operator $S : \mathbf{R}^3 \to \mathbf{R}^2$ gäller att $S(\mathbf{f}_3) = (\,-1 \quad 1\,)^T$ och att $S \circ T : \mathbf{R}^2 \to \mathbf{R}^2$ har matrisen

$$\begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$$

i den naturliga basen i $\mathbf{R}^2$. Bestäm matrisen för S i de naturliga baserna.

13. Antag att $T : E \to E$ är en linjär operator och att A och A' är matriser för T i olika baser. Visa att $\det(A) = \det(A')$.

14. Antag att S är speglingsmatrisen i ett hyperplan i $\mathbf{R}^n$ (dvs. i ett $(n-1)$-dimensionellt underrum) samt betrakta den linjära operatorn $T(\mathbf{x}) = S\mathbf{x}$. Visa med hjälp av uppgift 13 att $\det(S) = -1$. Ledning: Välj en ON-bas, i vilken matrisen för T är en diagonalmatris.

15. Betrakta operatorn $T : P_2 \to P_2$,

$$T(p)\,(x) = \int_0^\infty e^{-t} p(x-t)\,dt\,,$$

definierad på (t.ex.) vektorrummet P_2 av polynom av högst graden 2.

(a) Bestäm matrisen A för T i basen $\{1, x, x^2\}$ och dra av A:s utseende slutsatsen att T^{-1} existerar.

(b) Visa att $T^{-1} = Id + D$ (D är deriveringsoperatorn) **både** genom att invertera A och jämföra med matrisen för $Id + D$ **och** genom att använda partiell integration på uttrycket för $T(p')$.

Svar till övningsuppgifterna

Kapitel 1

1. $u = 3/2$, $v = -1/2$, $w = -3$.
2. $x_1 = \frac{1}{2}(-1 - s + t)$, $x_2 = 4 - t$, $x_3 = -1$, $x_4 = s$, $x_5 = 4 - t$, $x_6 = t$ ($x_4 = s$ och $x_6 = t$ är fria).
3. $$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ \phantom{x_1 + {}} x_2 + 2x_3 + 3x_4 = 1 \\ \phantom{x_1 + x_2 + {}} x_3 + x_4 = 1 \\ \phantom{x_1 + x_2 + x_3 + {}} x_4 = 1\,. \end{cases}$$
4. $x_1 = 19 + 2s$, $x_2 = -2 - 2s$, $x_3 = s$, $x_4 = s$ ($x_4 = s$ är en fri variabel).
5. $x_1 = -1$, $x_2 = 2$, $x_3 = 0$, $x_4 = 1$, $x_5 = -1$, $x_6 = -2$.
6. 22 000 fiskar i A, 1 000 fiskar i B.
7. I A finns dubbelt så många fiskar som i B.
8. Om x_1, x_2 och x_3 är antalet bakterier av typ 1, 2 respektive 3 som kan leva i tuben så gäller
$$x_2 = 15\,000 - 2s\,, \quad 0 \leq x_3 = s \leq 7\,500\,.$$
9. För $a \neq 3$, -1 är lösningen entydig:
$$x = -\frac{2}{a+1}\,, \quad y = \frac{a-1}{a+1}\,.$$
För $a = 3$ är $x = 1 - 3s$, $y = s$ ($y = s$ är fri). Det finns oändligt många lösningar. För $a = -1$ är systemet inkonsistent, dvs. lösningar saknas.
10. För $a = 0$ eller $a = -1$.
11. $a = 5$, $b = -3$, $c = -7$.
12. 7 000, 14 000 och 3 000 euro.

Kapitel 2

1. $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}$, $\quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
2. $\mathbf{x} = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}$.
3. $X = \begin{pmatrix} (5-5s)/2 & (3-5t)/2 \\ (-1+s)/2 & (1+3t)/6 \\ s & t \end{pmatrix}$, $\quad s$ och t är fria variabler.

4. $\begin{cases} x_{n+1} = 0,96x_n + 0,02y_n + 50 \\ y_{n+1} = 0,03x_n + 0,97y_n + 450 \end{cases}$, $\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 0,96 & 0,02 \\ 0,03 & 0,97 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} 50 \\ 450 \end{pmatrix}$.

5. $\mathcal{M} = \left\{ \begin{pmatrix} -s+t & s \\ s & t \end{pmatrix} \mid s, t \in \mathbf{R} \right\}$.

7. Exempel: Om A är en 3/2-matris så är A^T en 2/3-matris. Matriserna A^TA och AA^T är då av olika typ och kan inte vara lika. $A - A^T$ behöver inte vara symmetrisk.

11. $A = \frac{1}{2} \begin{pmatrix} 2 & -1 \\ -1 & 1 \\ 0 & 0 \end{pmatrix}$.

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

16. $a = 10/7$, $b = 9/7$, $c = 4/7$ och $d = 5/7$.

17. Proportionerna mellan mjölk, soja och vassla (mätt i g/hg) bör vara ungefär 272 : 395 : 235.

20. I t.ex. fall (c) får man, om man räknar upp grundämnena i ordningen Na, H, C, O, ekvationen

$$x_1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 8 \\ 6 \\ 7 \end{pmatrix} = x_3 \begin{pmatrix} 3 \\ 5 \\ 6 \\ 7 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

[Lösning: $(x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5)^T = (t/3)(3 \quad 1 \quad 1 \quad 3 \quad 3)^T$. Heltalslösningar fås t.ex. då $t = 3$.] I fallen (a) och (b) ställs ekvationen upp på liknanade sätt.

21. Förhållandet mellan priserna per tidsenhet för produktionen inom sektorerna (K), (E) och (M) bör vara 17 : 11 : 12.

23. (a) $T_1 = 45/2\,°C$, $T_2 = 30\,°C$, $T_3 = 20\,°C$ och $T_4 = 55/2\,°C$.
(b) $T_1 = T_4 = 120/7\,°C$, $T_2 = T_5 = 150/7\,°C$ och $T_3 = T_6 = 190/7\,°C$.

Kapitel 3

1. Om (j) och [j] betecknar rad j respektive kolonn j:

$$(a) \quad \begin{array}{l} E_{31}: \ (3) \to (3) + 4\,(1) \\ E_{23}: \ (2) \to (2) + 5\,(3) \end{array}, \qquad (b) \quad \begin{array}{l} E_{31}: \ [1] \to [1] + 4\,[3] \\ E_{23}: \ [3] \to [3] + 5\,[2] \end{array}.$$

2. $$\begin{cases} u - 2v + 3w = 11 \\ 4u + v - w = 4 \\ 2u - v + 3w = 10. \end{cases}$$

3. Ekvationen $A\mathbf{x} = \mathbf{b}$ kan skrivas $L(U\mathbf{x}) = \mathbf{b}$, ur vilket man löser ut $U\mathbf{x}$ och finner ett värde $\mathbf{c}$. Sedan löser man ut $\mathbf{x}$ ur $U\mathbf{x} = \mathbf{c}$. Lösningen blir $\mathbf{x} = (4 \quad 2 \quad 3)^T$.

4. Matriserna P_{ij} och P_{kl} kommuterar då $\{i,j\} = \{k,l\}$ samt då $\{i,j\} \cap \{k,l\} = \emptyset$.

5. (a) $3!$, (b) $n!$.

6. $$\begin{cases} (a) \quad L = \begin{pmatrix} 1 & & \\ -1/2 & 1 & \\ 0 & -2/3 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & -1 & 0 \\ & 3/2 & -1 \\ & & 4/3 \end{pmatrix}, \\ (b) \quad L = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ -1 & -1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & -1 \\ & -4 & 5 \\ & & 5 \end{pmatrix}. \end{cases}$$

7. Radbyte krävs om och endast om $a = 4$ och $b \neq 0$ (A är då icke-singulär). Matrisen A är singulär om och endast om $2ab - 6a - 3b + 24 = 0$.

8. $PA = LDU$ t.ex. då

$$(a) \quad P = P_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad (b) \text{ och } (c) \quad P = P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

varvid

$$(a) \quad L = \begin{pmatrix} 1 & \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & \\ & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 3/2 \\ & 1 \end{pmatrix};$$

$$(b) \quad L = \begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & & \\ & -1 & \\ & & -4 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & 3 \\ & 1 & 2 \\ & & 1 \end{pmatrix};$$

$$(c) \quad L = \begin{pmatrix} 1 & & \\ 3 & 1 & \\ 2 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & & \\ & -4 & \\ & & -5 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & 2 \\ & 1 & 3/2 \\ & & 1 \end{pmatrix}.$$

9. T.ex. för $P = P_{23}$, varvid

$$L = \begin{pmatrix} 1 & & & \\ 2 & 1 & & \\ -2 & 0 & 1 & \\ -1 & 1 & -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & -1 & 1 \\ & 1 & 3 & -3/2 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}.$$

11. $\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$.

12. (a) $\begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 2 \\ -1 & 1 & -1 \end{pmatrix}$. (b) Matrisen är singulär.

13. $C^{-1}BA^{-1}$. Ja. $A + B$ behöver inte vara inverterbar.

14. (a) $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, (b) $\frac{1}{4}\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

15. De 2/2-matriser som är sina egna inverser, är

$$\pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{och} \quad \begin{pmatrix} \pm\sqrt{1-bc} & b \\ c & \mp\sqrt{1-bc} \end{pmatrix}, \quad bc \leq 1.$$

16. $a = 3$, $a = 1/2$.

17. Vänsterinverserna till A är $\frac{1}{2}\begin{pmatrix} 1+s & 1-5s & 2s \\ -1+t & 1-5t & 2t \end{pmatrix}$, där s och t är fria.

18. $A = LDL^T$, där (a) $L = \begin{pmatrix} 1 & & \\ 2 & 1 & \\ 0 & 2 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}$;

(b) $L = \begin{pmatrix} 1 & & \\ 3 & 1 & \\ 5 & 1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & & \\ & 3 & \\ & & 2 \end{pmatrix}$.

24. (a) $A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

(b) $A = LU = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

25. $A = \frac{1}{42}\begin{pmatrix} -1 & 4 \\ 2 & -2 \end{pmatrix}$.

27. (a) LU-faktoriseringen är

$$\begin{pmatrix} 2 & 1 & & \\ 1 & 4 & 1 & \\ & 1 & 4 & 1 \\ & & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ 1/2 & 1 & & \\ & 2/7 & 1 & \\ & & 7/26 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & & \\ & 7/2 & 1 & \\ & & 26/7 & 1 \\ & & & 45/26 \end{pmatrix}.$$

(b) Den kubiska ri-funktionen är

$$f(x) = \begin{cases} 1 + (11t/4) - (3t^3/4), & \text{där } x = t \in [0,1] \\ 3 + (t/2) - (9t^2/4) + (3t^3/4), & \text{där } x = 1 + t \in [1,2]. \end{cases}$$

Kapitel 4

2. (a), (c) och (d).
3. (b) och (c).
5. $(\,t_1 \quad t_2\,) = (\,2 \quad 0\,)$.
7. (a) Linjärt beroende. (b) Linjärt beroende.
9. Snittet av de två underrummen är mängen av alla diagonalmatriser av typ n/n.
10. $R(A) = \operatorname{spn}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$, $N(A) = \operatorname{spn}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$, $R(B) = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}$, $N(B) = \mathbf{R}^3$.
11. (a) Echelonform för A är $U = \begin{pmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. I ekvationen $A\mathbf{x} = \mathbf{0}$, där $\mathbf{x} = (\,x_1 \quad x_2 \quad x_3 \quad x_4\,)^T$, är x_2 en basvariabel medan x_1, x_2 och x_4 är fria variabler.

$$N(A) = \operatorname{spn}\left\{(1 \quad 0 \quad 0 \quad 0)^T, (0 \quad -4 \quad 1 \quad 0)^T, (0 \quad 0 \quad 0 \quad 1)^T\right\}.$$

Ekvationen $A\mathbf{x} = \mathbf{b} = (\,b_1 \quad b_2\,)^T$ är konsistent om och endast en $2b_1 = b_2$ och härvid är lösningen

$$\mathbf{x} = b_1(0 \quad 1 \quad 0 \quad 0)^T + s(1 \quad 0 \quad 0 \quad 0)^T + t(0 \quad -4 \quad 1 \quad 0)^T + u(0 \quad 0 \quad 0 \quad 1)^T.$$

(b) Echelonform är $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}^T$. I ekvationen $A\mathbf{x} = \mathbf{0}$, där $\mathbf{x} = (\,x_1 \quad x_2\,)^T$, är x_1 basvariabel och x_2 fri variabel. $N(A) = \operatorname{spn}\left\{(\,-2 \quad 1\,)^T\right\}$. Ekvationen $A\mathbf{x} = \mathbf{b} = (\,b_1 \quad b_2 \quad b_3 \quad b_4\,)^T$ är konsistent om och endast om $b_1 = b_4 = 0$, $b_3 = 4b_2$ och då är lösningen $\mathbf{x} = b_2(\,1 \quad 0\,)^T + s(\,-2 \quad 1\,)^T$.

12. U och W är underrum, V och Z är inte underrum.
15. Linjärt beroende. En linjärkombination av vektorerna med t.ex. koefficienterna -1, 1, -1 och 1 blir noll.
16. (a) $\begin{cases} \text{Bas i } R(A)\text{:} \quad \{(1 \quad -1 \quad 5)^T, (-4 \quad 2 \quad -6)^T\}; \\ \text{Bas i } R(A^T)\text{:} \quad \{(1 \quad -4 \quad 9 \quad -7), (0 \quad -2 \quad 5 \quad -6)\}; \\ \text{Bas i } N(A)\text{:} \quad \{(2 \quad 5 \quad 2 \quad 0)^T, (-5 \quad -3 \quad 0 \quad 1)^T\}. \end{cases}$

    (b) $\begin{cases} \text{Bas i } R(A)\text{:} \quad \{(-2 \quad 1 \quad 3 \quad 1)^T, (-5 \quad 3 \quad 11 \quad 7)^T, (0 \quad 1 \quad 7 \quad 5)^T\}; \\ \text{Bas i } R(A^T)\text{:} \quad \{(1 \quad 3 \quad -5 \quad 1 \quad 5), (0 \quad 1 \quad -2 \quad 2 \quad -7), \\ \qquad\qquad\qquad\qquad (0 \quad 0 \quad 0 \quad -4 \quad 20)\}; \\ \text{Bas i } N(A)\text{:} \quad \{(-1 \quad 2 \quad 1 \quad 0 \quad 0)^T, (-1 \quad -3 \quad 0 \quad 5 \quad 1)^T\}. \end{cases}$
17. $\underline{a \neq 1}$: $\dim R(A) = 3$, $\dim N(A) = 1$.

$$\begin{aligned} &\text{Bas i } R(A): \quad \{(1 \quad 2 \quad 5)^T, (3 \quad 0 \quad 3)^T, (2 \quad a \quad 4)^T\}, \\ &\text{Bas i } N(A): \quad \{(1 \quad -1 \quad 0 \quad 2)^T\}, \end{aligned}$$

$\underline{a = 1}$: $\dim R(A) = 2$, $\dim N(A) = 2$.

$$\begin{aligned} &\text{Bas i } R(A): \quad \{(1 \quad 2 \quad 5)^T, (3 \quad 0 \quad 3)^T\}, \\ &\text{Bas i } N(A): \quad \{(-1 \quad -1 \quad 2 \quad 0)^T, (1 \quad -1 \quad 0 \quad 2)^T\}. \end{aligned}$$

18. $\begin{cases} (a) \text{ T.ex. } \{(1 \quad 0 \quad 1 \quad 0), (0 \quad 1 \quad 0 \quad 1), (-1 \quad 0 \quad 2 \quad 1). \\ (b) \text{ De fyra vektorerna är linjärt oberoende.} \\ (c) \text{ T.ex. } \{(1 \quad -1 \quad 1 \quad -1), (-1 \quad 1 \quad -1 \quad -1)\}. \end{cases}$
19. Nej.
21. $\begin{cases} (a) & \text{Linjärt oberoende. Bas i } \mathbf{R}^3. \\ (b) & \text{Linjärt beroende.} \\ (c) & \text{Linjärt oberoende. Ingen bas.} \\ (d) & \text{Linjärt beroende.} \\ (e) & \text{Linjärt beroende.} \end{cases}$
23. Linjärt oberoende.

24.
(a) En bas i underrummet är $\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&1\end{pmatrix}\right\}$. Dimensionen är 2.

(b) En bas i underrummet är $\left\{\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&2\end{pmatrix}\right\}$. Dimensionen är 2.
Koordinaterna för den givna matrisen är 2 och -3.

(c) En bas i underrummet är $\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&-1\\1&0\end{pmatrix}\right\}$. Dimensionen är 2.

25. Matrisen måste vara kvadratisk, dvs. av typen n/n.

26. (a) Falsk! T.ex. vektorerna $(1\ \ 0\ \ 0)$, $(0\ \ 1\ \ 0)$ och $(1\ \ 1\ \ 0)$ spänner upp xy-planet i $\mathbf{R}^3$, som ändå har dimensionen 2.
(b) Sann!
(c) Falsk! T.ex. är $\begin{pmatrix}1&1\\-1&-1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}1&1\\-1&-1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$. Implikationen stämmer bara om A är icke-singulär.

27. Underrummets dimension är 6 och **en** bas består av matriserna

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

31. (a) Koordinaterna är 0, -1 och 1.
(b) Koordinaterna för p i basen C är 1, -2 och 2. Koordinaterna för q i basen B är 8, 7 och 5.
(c) Polynomet $p(x)=3$ har de angivna koordinaterna.

32. $E=\begin{pmatrix}I&0\\-CA^{-1}&I\end{pmatrix}$, $E^{-1}=\begin{pmatrix}I&0\\CA^{-1}&I\end{pmatrix}$.

33. Inversen är $\begin{pmatrix}A^{-1}&0\\-D^{-1}CA^{-1}&D^{-1}\end{pmatrix}$.

34. Koordinaterna är x_1-x_2, x_2-x_3, x_3-x_4 och x_4.

35. De baser som uppfyller villkoret har formen $\left\{\frac{1}{2}\begin{pmatrix}3-3s\\7-3t\end{pmatrix},\begin{pmatrix}s\\t\end{pmatrix}\right\}$, där s och t uppfyller $7s\neq 3t$. **En** bas är t.ex. $\{\frac{1}{2}(0\ \ 7)^T,(1\ \ 0)^T\}$.

36. En bas för $V+W$ är $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{w}_1\}$ och en bas för $V\cap W$ är $\{(0\ \ -1\ \ 1\ \ 0)^T\}$.

39. Vi har t.ex. följande baser:

Bas i V: $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}=\{(1\ \ 0\ \ 0\ \ 0),(0\ \ -1\ \ 1\ \ 0),(0\ \ -1\ \ 0\ \ 1)\}$.
Bas i W: $\{\mathbf{w}_1,\mathbf{w}_2\}=\{(-1\ \ 1\ \ 0\ \ 0),(0\ \ 0\ \ 2\ \ 1)\}$.
Bas i $V\cap W$: $\{(3\ \ -3\ \ 2\ \ 1)\}$.
Bas i $V+W=\mathbf{R}^3$: $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{w}_1\}$.

Alltså är $\dim V=3$, $\dim W=2$, $\dim V\cap W=1$ och $\dim(V+W)=3$.

40. Mängden $\{\mathbf{v},\mathbf{w}\}$ är ingen bas i U. Däremot är t.ex. $\{\mathbf{v},\mathbf{w},(0\ \ 0\ \ -1\ \ 1)\}$ det.

41. En bas är t.ex. $\{(2\ \ 3\ \ 4),(0\ \ 1\ \ 1)\}$.

42. Ja, eftersom matrisens rang är 3, vilket också är antalet rader.
43. Det finns högerinverser och de är alla av formen $(1-s \quad s \quad t)^T$, där s och t är fria variabler. Vänsterinvers saknas.
44. $A = \mathbf{u}\mathbf{v}^T = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} (1 \quad -1)$ och $B = \mathbf{w}\mathbf{z}^T = \begin{pmatrix} 1 \\ 3 \end{pmatrix} (1 \quad 1 \quad 2)$.

Kapitel 5

1. (*a*) $\sqrt{61}$; (*b*) $t(21\,\mathbf{a} + \mathbf{b})$, där t är godtyckligt.
2. $a = 2 \pm \sqrt{3}$.
3. $2\pi/3$.
4. $\mathbf{v} = \pm\frac{1}{5}(3 \quad 4 \quad 0)^T$.
9. (*a*) $M_1 \perp M_2$ och $M_2 \perp M_4$; (*b*) M_2 och M_4 är varandras ortogonala komplement.
11. Alla vektorer av formen $t(1 \quad 1 \quad -2)$, $t \in \mathbf{R}$. För att ett ON-system skall uppstå, krävs bl.a. att $t \neq 0$.
14. Ja, $k = 5$ och $k = -4$. Mot $k = 5$ svarar $(a \quad b \quad c)^T = s(2 \quad 1 \quad 3)^T$ och mot $k = -4$ svarar $(a \quad b \quad c)^T = s(-5 \quad 2 \quad 6)^T$, $s \in \mathbf{R}$.
15. T.ex. matrisen $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.
16. (*a*) En bas i $W^\perp$ är

$$\{(-2 \quad 1 \quad 0 \quad 0 \quad 0)^T, (13 \quad 0 \quad -4 \quad 1 \quad 0)^T, (-17 \quad 0 \quad 5 \quad 0 \quad 1)^T\}.$$

(*b*) En bas i $U^\perp$ är $\{(-1 \quad 1 \quad 0)^T, (2 \quad 0 \quad 1)^T\}$ och $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, där $\mathbf{x}_1 = \frac{3}{2}(1 \quad 1 \quad -2)^T \in U$ och $\mathbf{x}_2 = -\frac{1}{2}(5 \quad 3 \quad 4)^T \in U^\perp$.
17. (*b*) $\mathbf{x} = (2 \quad 3 \quad 1 \quad 1)$.
19. Man kan välja $W = V^\perp = \operatorname{spn}\left\{(0 \quad 0 \quad 1 \quad 0)^T, (-2 \quad 1 \quad 0 \quad 1)^T\right\}$.

Kapitel 6

1. (*a*) -2 och 0. (*b*) $\det(A) = 20$.
3. (*a*) $\det(A) = 0$; (*b*) $\det(U) = 16$; (*c*) $\det(U^T) = 16$; (*d*) $\det(U^{-1}) = 1/16$; (*e*) $\det(M) = 16$.
4. (*b*) För alla 2/2-matriser A, för vilka $\det(A) = 0$.
5. $(x-y)(y-z)(z-x)$.
6. $(n-1)(-1)^{n-1}$.
8. (*a*) 12, (*b*) 39, (*c*) -36.
11. -1.
12. $A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.
13. $x = 3$, $y = -1$, $z = -2$.

17. $\det(A) = 0$ om $n \geq 3$, $\det(A) = -1$ då $n = 2$, $\det(A) = -1$ då $n = 1$.

18. (*a*) Matrisen A är inte inverterbar för $k = 1$ eller $k = -3$.

(*b*) Matrisen B är inte inverterbar för $k = 4$, $k = -2$ eller $k = 3$.

Kapitel 7

1. (*a*) $\lambda_1 = 9$, $\lambda_2 = -1$. Motsvarande egenvektorer är $s\,(3 \quad 7)^T$ respektive $s\,(-1 \quad 1)^T$ $(s \neq 0)$;

(*b*) $\lambda_1 = 3$, $\lambda_2 = -7$. Motsvarande egenvektorer är $s\,(3 \quad 1)^T$ respektive $s\,(-1 \quad 3)^T$ $(s \neq 0)$;

(*c*) $\lambda_1 = 5$ med egenvektorerna $s\,(1 \quad 1 \quad 1)^T$ $(s \neq 0)$ samt $\lambda_2 = 2$ med egenvektorerna $s\,(-1 \quad 1 \quad 0)^T + t\,(-1 \quad 0 \quad 1)^T$ ($s \neq 0$ eller $t \neq 0$).

2. (*a*) $\begin{cases} \lambda_1 = 12. & \text{Bas i } V(12)\text{: } \{(3 \quad 5)^T\}. \\ \lambda_2 = -4. & \text{Bas i } V(-4)\text{: } \{(-1 \quad 1)^T\}. \end{cases}$

(*b*) $\begin{cases} \lambda_1 = 10. & \text{Bas i } V(10)\text{: } \{(-1 \quad 2)^T\}. \\ \lambda_2 = 5. & \text{Bas i } V(5)\text{: } \{(2 \quad 1)^T\}. \end{cases}$

(*c*) $\begin{cases} \lambda_1 = 5. & \text{Bas i } V(5)\text{: } \{(0 \quad -1 \quad 1)^T\}. \\ \lambda_2 = 3. & \text{Bas i } V(3)\text{: } \{(0 \quad 1 \quad 0)^T, (-1 \quad 0 \quad 1)^T\}. \end{cases}$

(*d*) $\begin{cases} \lambda_1 = 3. & \text{Bas i } V(3)\text{: } \{(0 \quad 1 \quad 2)^T\}. \\ \lambda_2 = 2. & \text{Bas i } V(2)\text{: } \{(2 \quad 1 \quad 1)^T\}. \\ \lambda_3 = -1. & \text{Bas i } V(-1)\text{: } \{(-1 \quad 0 \quad 1)^T\}. \end{cases}$

3. (*a*) $\begin{cases} \lambda_1 = 2^{62}. & \text{Bas i } V(2^{62})\text{: } \{(-1 \quad 1 \quad 0)^T\}. \\ \lambda_2 = 1. & \text{Bas i } V(1)\text{: } \{(1 \quad -1 \quad 1)^T, (0 \quad -1 \quad 1)^T\}. \end{cases}$

(*b*) $\begin{cases} \lambda_1 = 4^{62}. & \text{Bas i } V(4^{62})\text{: } \{(0 \quad 1 \quad 1)^T\}. \\ \lambda_2 = 2^{62}. & \text{Bas i } V(2^{62})\text{: } \{(1 \quad 1 \quad 1)^T, (1 \quad 1 \quad 0)^T\}. \end{cases}$

5. $\mathbf{x} \perp \mathbf{y}$.

Kapitel 8

1. En ON-bas i planet är t.ex. $\{\frac{1}{\sqrt{2}}(1 \quad 1 \quad 0)^T, \frac{1}{\sqrt{6}}(-1 \quad 1 \quad 2)^T\}$. Projektionsmatrisen är

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

2. (*a*) $\mathbf{p}_1 = \frac{10}{3}(1 \quad 1 \quad 1)^T$. (*b*) $\mathbf{p}_2 = \frac{5}{9}(1 \quad 2 \quad 2)^T$.

3. (*a*) $\mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $\mathbf{p} = \mathbf{b} = (1 \quad 3 \quad 4)^T$ ($\mathbf{b}$ ligger i A:s kolonnrum!).

(*b*) $\mathbf{x} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{p} = \frac{4}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. (*c*) $\mathbf{x} = \begin{pmatrix} 16/3 \\ -3/2 \end{pmatrix}$, $\mathbf{p} = \frac{1}{6}\begin{pmatrix} 23 \\ 14 \\ 5 \end{pmatrix}$.

5. (*a*) $\mathbf{p}_1 = \frac{1}{3}(4 \quad 4 \quad -2)^T$, $\mathbf{p}_2 = \frac{1}{3}(-2 \quad 4 \quad 4)^T$. Projektionen på $\operatorname{spn}\{\mathbf{a}_1, \mathbf{a}_2\}$ är $\mathbf{p}_1 + \mathbf{p}_2 = \frac{2}{3}(1 \quad 4 \quad 1)^T$.

(*b*) $\mathbf{p}_3 = -\frac{1}{3}(2 \quad -1 \quad 2)^T$.

6. Bas för $N(A)$ är $\{(2 \quad 2 \quad -1)^T\}$. Vidare är $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, där $\mathbf{x}_1 = (1 \quad 1 \quad 4)^T$ och $\mathbf{x}_2 = (2 \quad 2 \quad -1)^T$. Koordinaterna för $\mathbf{x}_1$ i den angivna basen är 0 och 1.

7. (*a*) $A = 1\,P_1 + 5\,P_2 = 1 \cdot \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + 5 \cdot \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

$$A^n = P_1 + 5^n P_2 = \frac{1}{2}\begin{pmatrix} 5^n+1 & 5^n-1 \\ 5^n-1 & 5^n+1 \end{pmatrix}. \quad A^{1/3} = \frac{1}{2}\begin{pmatrix} \sqrt[3]{5}+1 & \sqrt[3]{5}-1 \\ \sqrt[3]{5}-1 & \sqrt[3]{5}+1 \end{pmatrix}.$$

(*b*) $A = 1\,P_1 + (-2)P_2 = 1 \cdot \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} + (-2) \cdot \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

$$A^n = P_1 + (-2)^n P_2 = \frac{1}{3}\begin{pmatrix} 2+(-2)^n & -1+(-2)^n & -1+(-2)^n \\ -1+(-2)^n & 2+(-2)^n & -1+(-2)^n \\ -1+(-2)^n & -1+(-2)^n & 2+(-2)^n \end{pmatrix}.$$

$$A^{1/3} = \frac{1}{3}\begin{pmatrix} 2-\sqrt[3]{2} & -1-\sqrt[3]{2} & -1-\sqrt[3]{2} \\ -1-\sqrt[3]{2} & 2-\sqrt[3]{2} & -1-\sqrt[3]{2} \\ -1-\sqrt[3]{2} & -1-\sqrt[3]{2} & 2-\sqrt[3]{2} \end{pmatrix}.$$

9. Avståndet är $|\mathbf{n}^T\mathbf{y} - b| \,/\, \|\mathbf{n}\|$.

10. (*a*) $A = QR$, där $Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ och $R = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

(*b*) $A = QR$, där $Q = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$ och $R = \begin{pmatrix} 3 & -6 & -7/3 \\ 0 & 3 & 14/3 \\ 0 & 0 & 5/3 \end{pmatrix}$.

11. (*a*) *ON*-bas: $\left\{\frac{1}{\sqrt{2}}(1 \quad 1 \quad 0)^T, \frac{1}{\sqrt{3}}(-1 \quad 1 \quad 1)^T\right\}$.

(*b*) Projektionsmatrisen är $P = \frac{1}{6}\begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix}$.

(*c*) Speglingen av $\mathbf{x}$ är vektorn $S\mathbf{x} = \frac{1}{3}\begin{pmatrix} 2x+y-2z \\ x+2y+2z \\ -2x+2y-z \end{pmatrix}$.

13. Volymen är 27.

Kapitel 9

1. (*a*) Diagonaliserbar. (*b*) och (*c*) Inte diagonaliserbara.

2. $B^{-1}AB = D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, om $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Härvid är

$A^n = BD^nB^{-1} = \begin{pmatrix} 1 & 0 \\ 1-2^n & 2^n \end{pmatrix}$. En kvadratrot är $\begin{pmatrix} 1 & 0 \\ 1-\sqrt{2} & \sqrt{2} \end{pmatrix}$.

3. $Q^TAQ = D = \begin{pmatrix} 0 & & \\ & 3 & \\ & & 3 \end{pmatrix}$, om $Q = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{pmatrix}$.

4. $Q^TAQ = D = \begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}$, om $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

Härvid är $A^n = QD^nQ^T = \frac{1}{2}\begin{pmatrix} (a+b)^n + (a-b)^n & (a+b)^n - (a-b)^n \\ (a+b)^n - (a-b)^n & (a+b)^n + (a-b)^n \end{pmatrix}$.

5. $A = \frac{1}{3}\begin{pmatrix} -4 & -5 \\ -10 & 1 \end{pmatrix}$.

6. (a) $\quad A = \begin{pmatrix} b_0 & b_1 & \cdots & b_{n-1} & b_n \\ f_0 & 0 & \cdots & 0 & 0 \\ 0 & f_1 & \cdots & 0 & 0 \\ - & - & - & - & - \\ 0 & 0 & \cdots & f_{n-1} & 0 \end{pmatrix}$.

(c) A har egenvärdet 1 samt två egenvärden, vilkas belopp ligger mellan 0 och 1. En konstant åldersfördelning är alltså möjlig. En liten störning i betingelserna kan emellertid göra att alla egenvärden blir till beloppet mindre än 1. Om detta händer riskerar arten dö ut.

7. Mor Stava är glad 2/3 av sina levnadsdagar.

Kapitel 10

3. (a) Ja. $\quad (b)$ Nej. $\quad (c)$ Ja.

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

5. $\mathbf{x} = t\,(3 \quad 3 \quad -1)^T, \quad t \in \mathbf{R}$.

6. (b) $\begin{pmatrix} 9/2 & 1 & 1/2 \\ 4 & 1 & 4 \\ -1/2 & 1 & 3/2 \end{pmatrix}$.

7. $\mathbf{b}'_1 = 3\mathbf{b}_1 + \mathbf{b}_2, \quad \mathbf{b}'_2 = -\mathbf{b}_1 + 2\mathbf{b}_2$.

9. $\begin{cases} (a) \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \qquad (b) \quad \begin{pmatrix} 0 & 1 & -2 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \\ (c) \quad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \qquad (d) \quad \begin{pmatrix} 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{pmatrix}. \end{cases}$

10. $\begin{cases} x_1 = x'_1 + x'_2 \\ x_2 = 2x'_1 + x'_2 + x'_3 \\ x_3 = -3x'_1 + x'_3\,. \end{cases}$

11. Koordinaterna 3, 1 och 0. Koordinaterna 9, -14 och 7.

12. $\begin{pmatrix} 2 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$.

15. (a) $\quad A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.