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 $t_0 = \mathbf{a}^T\mathbf{b}/\|\mathbf{a}\|^2$.

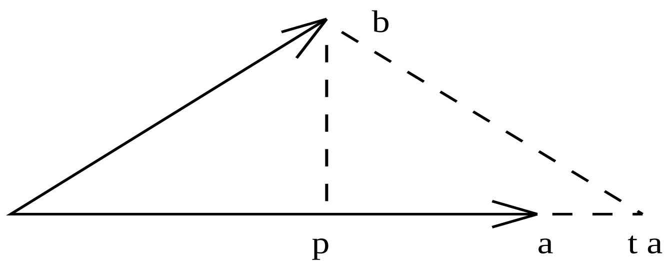

fig. 1

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

$$\mathbf{p} = t_0\mathbf{a} = \frac{\mathbf{a}^T\mathbf{b}}{\|\mathbf{a}\|^2}\mathbf{a}\,, \tag{1}$$

som vi också kan skriva

$$\mathbf{p} = \frac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2}\mathbf{b}\,, \tag{2}$$

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 $A\mathbf{x} = \mathbf{b}$.

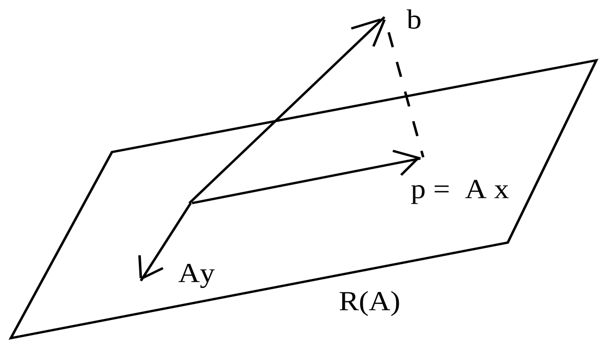

fig. 2

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 $A = (\,\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n\,)$ 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 $A^T A$ inverterbar. Då är $\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$ 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 $A^T A$ 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 $\mathbf{x} \in N(A^T A) = N(A)$ 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 $(A^T A \mid A^T\mathbf{b})$. 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

$$P = A(A^T A)^{-1} A^T \tag{3}$$

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 $R(A) = \mathrm{spn}\,\{\mathbf{a}\}$. 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

$$\begin{aligned} P^T = (A(A^TA)^{-1}A^T)^T &= A^{TT}((A^TA)^{-1})^T A^T \\ &= A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T = P\,, \end{aligned}$$

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 $R(P)$.

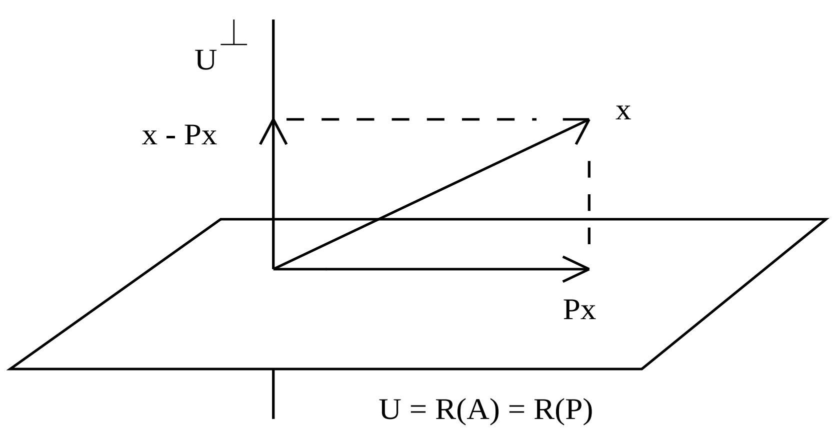

fig. 3

(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:

$$A^TA = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_p \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \ldots & \mathbf{a}_1^T\mathbf{a}_p \\ \vdots & & \vdots \\ \mathbf{a}_p^T\mathbf{a}_1 & \ldots & \mathbf{a}_p^T\mathbf{a}_p \end{pmatrix} = I_p\,. \tag{5}$$

För projektionsmatrisen $P = A(A^TA)^{-1}A^T$ på $R(A)$ får vi därför uttrycket

$$P = AA^T\,. \tag{6}$$

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:

$$P = AA^T = \begin{pmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_p \end{pmatrix} \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p \end{pmatrix} = \mathbf{a}_1\mathbf{a}_1^T + \cdots + \mathbf{a}_p\mathbf{a}_p^T\,. \tag{7}$$

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.

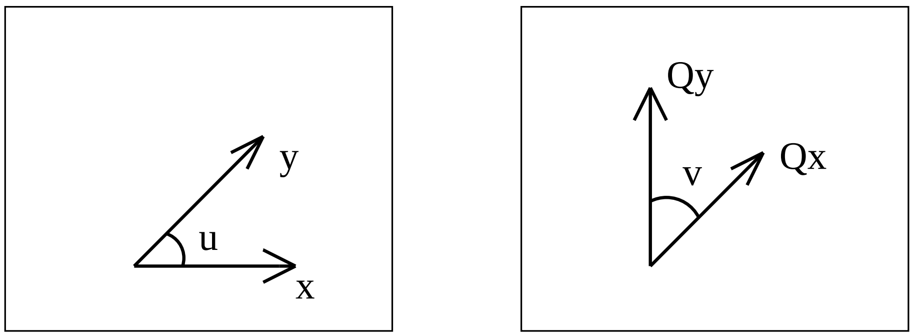

fig. 4

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

$$S\mathbf{x} = x + 2(P\mathbf{x} - \mathbf{x}) = 2P\mathbf{x} - \mathbf{x} = (2P - I)\mathbf{x}$$

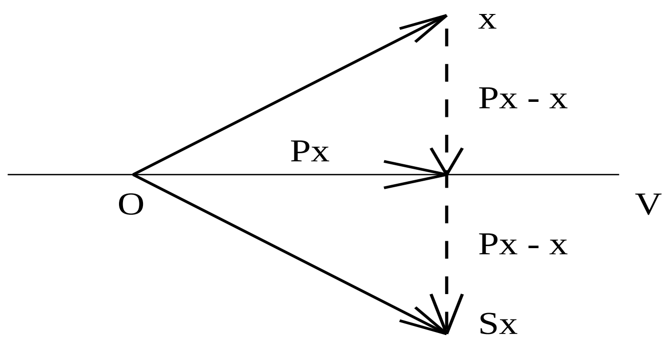

fig. 5

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 $S^T S = S^2 = I$, 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 $\mathbf{q}_1, \ldots, \mathbf{q}_p$ bildar nu ett ON-system med samma spann som de ursprungliga vektorerna $\mathbf{a}_1, \ldots, \mathbf{a}_p$.

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

$$0 = \mathbf{v}_1^T\mathbf{v}_3 = \mathbf{v}_1^T\mathbf{a}_3 - \lambda_1\|\mathbf{v}_1\|^2 = 1 - 2\lambda_1\,,$$
$$0 = \mathbf{v}_2^T\mathbf{v}_3 = \mathbf{v}_2^T\mathbf{a}_3 - \lambda_2\|\mathbf{v}_2\|^2 = \frac{1}{2} - \frac{3}{2}\lambda_2\,.$$

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.

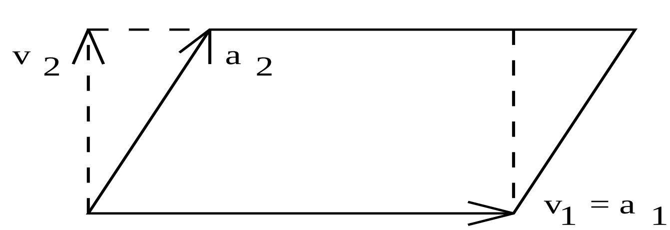

fig. 6

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 $\operatorname{spn}\{\mathbf{b}\}$, 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 $\operatorname{spn}\{\mathbf{a}_1, \mathbf{a}_2\}$.

   (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

$$(a) \quad \mathbf{a}_1 = (\,0 \quad 0 \quad 1\,)^T\,, \quad \mathbf{a}_2 = (\,0 \quad 1 \quad 1\,)^T\,, \quad \mathbf{a}_3 = (\,1 \quad 1 \quad 1\,)^T\,;$$
$$(b) \quad \mathbf{a}_1 = (\,-1 \quad 2 \quad 2\,)^T\,, \quad \mathbf{a}_2 = (\,4 \quad -5 \quad -2\,)^T\,, \quad \mathbf{a}_3 = (\,5 \quad -2 \quad 1\,)^T\,.$$

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.