+AV

1. Otetaan 2x2 matriisi $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Muodostetaan karakteristinen polynomi :

$$\det(A - \lambda I) = 0$$

$$\Leftrightarrow \quad \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow \quad (a-\lambda)(d-\lambda) - bc = 0$$

$$\Leftrightarrow \quad \lambda^2 - (a+d)\lambda + ad - bc = 0$$

Huomataan : $tr(A) = a+d$, $\det(A) = ad - bc$

2. asteen polynomilla on reaaliset juuret, jos ns. diskriminantti $D \geq 0$ (Jos $D = 0$, vain yksi juuri)

nyt $D = tr(A)^2 - 4\det(A) \geq 0$

$$\Leftrightarrow \quad tr(A)^2 \geq 4\det(A)$$

(Toisen asteen ratkaisukaava lienee muistissa :))

2. a) $A = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

Kar. pol. : $\begin{vmatrix} 1-\lambda & -2 \\ 1 & 3-\lambda \end{vmatrix} = 0$

$(1-\lambda)(3-\lambda)+2 = 0$

$3-\lambda-3\lambda+\lambda^2+2 = 0$

$\lambda^2 - 4\lambda + 5 = 0$

$\lambda = \frac{4 \pm \sqrt{16-20}}{2} = 2 \pm \frac{\sqrt{4i^2}}{2} = \begin{cases} 2+i \\ 2-i \end{cases}$

Ratkaistaan ominaisvektori :

$\lambda = 2+i$: $(A-\lambda I) = \begin{bmatrix} -1-i & -2 \\ 1 & 1-i \end{bmatrix} \xrightarrow{\cdot\frac{(1-i)}{2}} \sim \begin{bmatrix} -1-i & -2 \\ 0 & 0 \end{bmatrix}$

$\Rightarrow (-1-i)x_1 - 2x_2 = 0$

valitaan $x_1 = 1 \Rightarrow x_2 = -\frac{1}{2} - \frac{1}{2}i$ eli $v_1 = \begin{bmatrix} 1 \\ -\frac{1}{2}-\frac{1}{2}i \end{bmatrix}$

Toista ominaisarvoa vastaava vektori on edellisen kompleksi konjugaatti (komponentit ovat kompl. konj. !)

Muodostetaan ominaisvektoreista matriisi V :

$V = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2}-\frac{1}{2}i & -\frac{1}{2}+\frac{1}{2}i \end{bmatrix}$ Lasketaan V^{-1} Cramerin säännöllä :

$\det(V) = -\frac{1}{2}+\frac{1}{2}i+\frac{1}{2}+\frac{1}{2}i = i$

$\frac{1}{i} = -i$

$V^{-1} = \frac{1}{i}\begin{bmatrix} -\frac{1}{2}+\frac{1}{2}i & -1 \\ \frac{1}{2}+\frac{1}{2}i & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2i}+\frac{1}{2} & -\frac{1}{i} \\ \frac{1}{2i}+\frac{1}{2} & \frac{1}{i} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{1}{2}i & i \\ \frac{1}{2}-\frac{1}{2}i & -i \end{bmatrix}$

A:n kompleksinen diagonalisointi

$A = VDV^{-1} = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2}-\frac{1}{2}i & -\frac{1}{2}+\frac{1}{2}i \end{bmatrix}\begin{bmatrix} 2+i & 0 \\ 0 & 2-i \end{bmatrix}\begin{bmatrix} \frac{1}{2}+\frac{1}{2}i & i \\ \frac{1}{2}-\frac{1}{2}i & -i \end{bmatrix} \left(= \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix} \right)$

Todellakin!

2. b) Ominaisarvo $\lambda = 2 + i$

Kiertomatriisi $C = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$

$P = [\operatorname{Re} v \mid \operatorname{Im} v] = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$, käänteismatriisi Cramerin säännöllä:

$$P^{-1} = \frac{1}{(-\frac{1}{2})} \begin{bmatrix} -\frac{1}{2} & 0 \\ \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix}$$

$$A = PCP^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix}$$

Jos käytetään ominaisarvoa $\lambda = 2 - i$ ja siihen liittyvää ominaisvektoria, saadaan:

$$C = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$$

3. $V_1 = [3, 1, 1]$ $\quad V_2 = [-1, 2, 1]$ $\quad V_3 = [-\frac{1}{2}, -2, \frac{7}{2}]$

Vektorit ovat ortogonaaliset jos niiden pistetulo $= 0$;

$$V_1 \cdot V_2 = [3, 1, 1]\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = -3 + 2 + 1 = 0$$

$$V_1 \cdot V_3 = -\frac{3}{2} - 2 + \frac{7}{2} = 0 \quad \text{ja} \quad V_2 \cdot V_3 = \frac{1}{2} - 4 + \frac{7}{2} = 0$$

$\Rightarrow$ vektorit ovat keskenään ortogonaalisia.

Halutaan ilmoittaa vektori $V = [6, 1, 8]$ tässä kannassa ts. kantavektoreiden lineaarikombinaationa :

$$V = C_1 V_1 + C_2 V_2 + C_3 V_3$$

| Lasketaan pistetulo V_1:n kanssa

$$\Rightarrow V \cdot V_1 = C_1 V_1 \cdot V_1 + C_2 \underbrace{V_2 \cdot V_1}_{=0} + C_3 \underbrace{V_3 \cdot V_1}_{=0}$$

$$\Rightarrow C_1 = \frac{V \cdot V_1}{\|V_1\|^2}, \quad V_1 \cdot V_1 = [3, 1, 1]\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = 9 + 1 + 1 = 11$$

$$V \cdot V_1 = [6\ 1\ 8]\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = 18 + 1 + 8 = 27$$

$\Rightarrow C_1 = \frac{27}{11}$, Vastaavasti lasketaan C_2 ja C_3.

$$C_2 = \frac{V \cdot V_2}{\|V_2\|^2} = \dots = \frac{2}{3} \qquad C_3 = \frac{V \cdot V_3}{\|V_3\|^2} = \dots = \frac{46}{33}$$

$$V = \frac{27}{11} V_1 + \frac{2}{3} V_2 + \frac{46}{33} V_3$$

4.

$$A = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & g \end{bmatrix} \qquad A_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

A:n ominaisarvot:

$\det(A - \lambda I) = 0$

$$\begin{vmatrix} a-\lambda & b & 0 \\ c & d-\lambda & 0 \\ 0 & 0 & g-\lambda \end{vmatrix} = \underset{\text{ominaisarvo } g}{\underset{\uparrow}{(g-\lambda)}} \underbrace{\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix}}_{A_1\text{:n ominaisarvot}} = 0$$

Ominaisvektorit:

Jos sijoitetaan A_1:n ominaisarvot, λ,

$$(A - \lambda I)x = \begin{bmatrix} a-\lambda & b & 0 \\ c & d-\lambda & 0 \\ 0 & 0 & g-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \bar{0}$$

nähdään että $x_3 = 0$ ja $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ saadaan A_1:n ominaisvektoreista (annettu v_1 ja v_2)

Kun sijoitetaan $\lambda = g \Rightarrow x_3$ on vapaasti valittavissa, valitaan $x_3 = 1$

A:n kaikki ominaisvektorit ovat siis $\begin{bmatrix} v_1 \\ 0 \end{bmatrix}, \begin{bmatrix} v_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Huom. Tällainen lohkodiagonaalimatriisi kuvaa siis vektorien vastaavat lohkot toisistaan riippumatta, mikä näkyy ominaisarvoissa ja -vektoreissa.

5.

a)

$$A = \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{bmatrix} \end{array}$$

← merkitään sarakkeet

Sarakkeiden väliset pistetulot :

$$a_1 \cdot a_2 = -ab + ab + 0 = 0$$
$$a_1 \cdot a_3 = a \cdot 0 + b \cdot 0 + 0 \cdot c = 0$$
$$a_2 \cdot a_3 = -b \cdot 0 + a \cdot 0 + 0 \cdot c = 0$$

Ovat ortogonaaliset !

Sarakkeet on normeerattava tekijällä $\sqrt{a_i \cdot a_i}$, $i = 1, 2, 3$
(eli vektoreiden pituuksilla $\|a_i\|$)

Sarake 1 : $\sqrt{a_1 \cdot a_1} = \sqrt{a^2 + b^2}$

Sarake 2 : $\sqrt{a_2 \cdot a_2} = \sqrt{b^2 + a^2}$

Sarake 3 : $\sqrt{a_3 \cdot a_3} = \sqrt{c^2} = c$

(Huom! On siis jaettava näillä "kertoimilla")

b)

$$A = \begin{bmatrix} 0.8 & -0.6 & 0 \\ 0.6 & 0.8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$$

Aloitetaan määrittelemällä alimatriisin $A_1 = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$ ominaisarvot

$$\det(A_1 - \lambda I) = \begin{vmatrix} 0.8-\lambda & -0.6 \\ 0.6 & 0.8-\lambda \end{vmatrix} = (0.8-\lambda)^2 + 0.36 = 0$$

$$\Leftrightarrow \lambda = \begin{cases} 0.8 - 0.6i \\ 0.8 + 0.6i \end{cases}$$

Ratkaistaan ominaisvektori, kun $\lambda = 0.8 - 0.6i$

$$(A_1 - \lambda I) = \begin{bmatrix} 0.6i & -0.6 \\ 0.6 & 0.6i \end{bmatrix} \sim \begin{bmatrix} 0.6i & -0.6 \\ 0 & 0 \end{bmatrix} \Rightarrow i x_1 = x_2$$

valitaan $x_1 = 1$

$$\Rightarrow v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Kuten tehtävässä 2, toinen ominaisvektori saadaan suoraan : $v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$

(5.) jatkuu...

kuten tehtävä 4 osoitti, A:n ominaisarvot ovat alimatriisin A_1 ominaisarvot ja 1.07. Ominaisvektorit saadaan A_1:n ominaisvektoreista, ja $\lambda = 1.07$ vastaava vektori $= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(Sij. $\lambda = 1.07$: $(A - \lambda I) = \begin{bmatrix} -0.27 & -0.6 & 0 \\ 0.6 & -0.27 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$X_1 = 0$, $X_2 = 0$, X_3 vapaa, valitaan $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$)

A:n ominaisarvot $\lambda = \begin{cases} 0.8 - 0.6i \\ 0.8 + 0.6i \\ 1.07 \end{cases}$

A:n ominaisvektorit $V_1 = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}$ $V_2 = \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}$ $V_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(6.) $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

Tiedetään, että yksi ominaisarvo on $\lambda_1 = 5$ ja yksi ominaisvektori $w_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

Lasketaan w_2:ta vastaava ominaisarvo λ_2:

$$A w_2 = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} = 2 w_2$$

Siis $\lambda_2 = 2$.

Kolmas ominaisarvo voidaan selvittää kahdella tavalla:

1° Muodostetaan karakteristinen polynomi $\det(A - \lambda I)$ ja jaetaan se tekijöillä $\lambda - 5$ ja $\lambda - 2$.
(SUORAVIIVAINEN, MUTTA PITKÄHKÖ TAPA)

2° Huomataan, että jos 3. asteen polynomilla on juuret λ_1, λ_2 ja λ_3, niin se on muotoa

$$a(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$$
$$= a\lambda^3 - a(\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 + a(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)\lambda - a\lambda_1\lambda_2\lambda_3$$

jolloin juurien tulo $\lambda_1\lambda_2\lambda_3$ on yhtä suuri kuin

6... (5.2)

vakiotermin kerroin jaettuna kolmannen asteen kertoimen vastaluvulla $-a$.

Karakteristisen polynomin $\det(A-\lambda I)$ vakiotermi on

$$\det(A-0I) = \begin{vmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{vmatrix} = 20$$

ja sen kolmannen asteen termin kerroin on $a=-1$.

Siis

$$\lambda_1\lambda_2\lambda_3 = \frac{20}{-(-1)} = 20 \quad \Rightarrow \lambda_3 = \frac{20}{2\cdot 5} = 2.$$

Ominaisarvot ovat siis $\lambda_1 = 5$ ja $\lambda_2 = \lambda_3 = 2$.

Ortogonaalinen diagonalisointi: kirjoitetaan

$$A = VDV^{-1},$$

missä

$$V = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \qquad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

ja v_1, v_2 ja v_3 ovat ortonormaalit λ_1:tä, λ_2:ta ja λ_3:a vastaavat ominaisvektorit.

6... (s.3) Lasketaan ensin jotkin ominaisvektorit

w_1, w_2 ja w_3 ja ortonormeerataan ne tarvittaessa:

$$Aw_1 = 5w_1 : \left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right]$$

JÄTETÄÄN OIKEA PUOLI KIRJOITTAMATTA ALLA, KOSKA NOLLAVEKTORI SÄILYY RIVI-OPERAATIOISSA NOLLANA.

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{array}{cc} 2 & -1 \end{array} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix} \begin{array}{c} 1 \end{array} \cdot\left(-\tfrac{1}{3}\right) \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

x_3 vapaa; $:= t$, $x_2 = x_3 = t$, $x_1 = 2x_2 - x_3 = t$

Siis $w_1 = t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$w_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ tiedettiinkin jo

$$Aw_3 = 2w_3 : \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{array}{cc} -1 & -1 \end{array} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

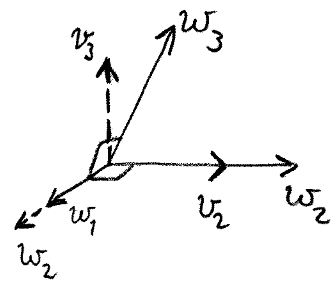
x_3 vapaa, $:= t$ x_2 vapaa, $:= s$, $x_1 = -x_2 - x_3 = -s - t$

Valinta $s = 1$, $t = 0$ antaa w_2:n.

Valitaan w_3:ta varten esim. $s = 0$, $t = 1$ $\rightsquigarrow$ $w_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

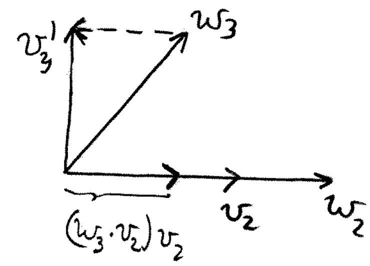
6... (5.4) Ortonormeeraus:

Koska A on symmetrinen, sen erisuuria ominaisarvoja vastaavat ominaisvektorit ovat valmiiksi kohtisuorassa toisiaan vastaan: $w_1 \perp w_2$, $w_1 \perp w_3$ (Ortogonaaliset...-pruju, lause 1.6)
Siis riittää normeerata $w_1 \rightsquigarrow v_1$ ja ortonormeerata $w_2 \rightsquigarrow v_2$ ja $w_3 \rightsquigarrow v_3$ keskenään:



$$v_1 = \frac{1}{\|w_1\|} w_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Gram-Schmidt: $(w_1, w_2) \rightsquigarrow (v_1, v_2)$



$$v_2 = \frac{1}{\|w_2\|} w_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$v_3' = w_3 - (w_3 \cdot v_2) v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}$$

$$v_3 = \frac{1}{\|v_3'\|} v_3 = \frac{\sqrt{2}}{\sqrt{3}} \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

6... (5.5)

Muodostetaan ortonormaaleista vektoreista matriisi U, jonka käänteismatriisi saadaan helposti transponoimalla: $U^{-1} = U^T$:

$$U = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & \sqrt{2}/\sqrt{3} \end{bmatrix} \approx \begin{bmatrix} 0{,}5774 & -0{,}7071 & -0{,}4082 \\ 0{,}5774 & 0{,}7071 & -0{,}4082 \\ 0{,}5774 & 0 & 0{,}8165 \end{bmatrix}$$

$$D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$U^{-1} = U^T = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & \sqrt{2}/\sqrt{3} \end{bmatrix}$$

$$A = U D U^T$$

(Helppo tarkistaa esim. Matlabilla, samoin kuin se, että $U^{-1} = U^T$.)

Liitteenä muutama sivu kirjasta Golubitsky, Dellnitz: Linear Algebra and Differential Equations Using Matlab

Note that if $u, v \in \mathbf{R}^n$ are column vectors, then $u \cdot v = u'v$. Therefore we can rewrite (10.2.5) as

$$A'A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = A'x_0,$$

where A is the matrix whose columns are the w_j and x_0 is viewed as a column vector. Note that the matrix $A'A$ is a $k \times k$ matrix.

We claim that $A'A$ is invertible. To verify this claim, it suffices to show that the null space of $A'A$ is 0; that is, if $A'Az = 0$ for some $z \in \mathbf{R}^k$, then $z = 0$. First, calculate

$$||Az||^2 = Az \cdot Az = (Az)'Az = z'A'Az = z'0 = 0.$$

It follows that $Az = 0$. Second, if we let $z = (z_1, \ldots, z_k)'$, then the equation $Az = 0$ may be rewritten as

$$z_1 w_1 + \cdots + z_k w_k = 0.$$

Since the w_j are linearly independent, it follows that the $z_j = 0$. In particular, $z = 0$. Since $A'A$ is invertible, (10.2.4) is valid and the theorem is proved. ◆

Gram-Schmidt Orthonormalization Process

Suppose that $\mathcal{W} = \{w_1, \ldots, w_k\}$ is a basis for the subspace $V \subset \mathbf{R}^n$. There is a natural process by which the $\mathcal{W}$ basis can be transformed into an orthonormal basis $\mathcal{V}$ of V. This process proceeds inductively on the w_j; the orthonormal vectors $v_1, \ldots, v_k$ can be chosen so that

$$\operatorname{span}\{v_1, \ldots, v_j\} = \operatorname{span}\{w_1, \ldots, w_j\}$$

for each $j \leq k$. Moreover, the v_j are chosen using the theory of least squares that we have just discussed.

The Case $j = 2$

To gain a feeling for how the induction process works, we verify the case $j = 2$. Set

$$v_1 = \frac{1}{||w_1||} w_1, \tag{10.2.6}$$

so v_1 points in the same direction as w_1 and has unit length—that is, $v_1 \cdot v_1 = 1$. The normalization is shown in Figure 10.2.

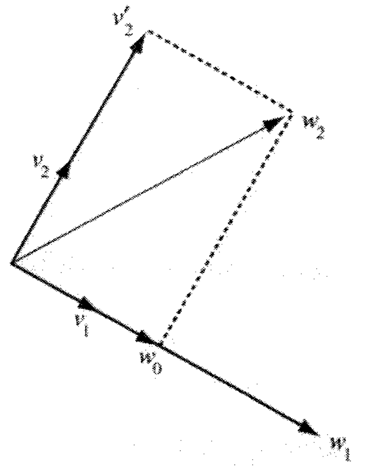


Figure 10.2
Planar illustration of Gram-Schmidt orthonormalization

Next we find a unit length vector v_2' in the plane spanned by w_1 and w_2 that is perpendicular to v_1. Let w_0 be the vector on the line generated by v_1 that is nearest to w_2. It follows from (10.2.3) that

$$w_0 = \frac{w_2 \cdot v_1}{||v_1||^2} v_1 = (w_2 \cdot v_1) v_1.$$

The vector w_0 is shown on Figure 10.2 and, as Lemma 10.2.1 states, the vector $v_2' = w_2 - w_0$ is perpendicular to v_1. That is,

$$v_2' = w_2 - (w_2 \cdot v_1) v_1 \qquad \textbf{(10.2.7)}$$

is orthogonal to v_1.

Finally, set

$$v_2 = \frac{1}{||v_2'||} v_2' \qquad \textbf{(10.2.8)}$$

so that v_2 has unit length. Since v_2 and v_2' point in the same direction, v_1 and v_2 are orthogonal. Note also that v_1 and v_2 are linear combinations of w_1 and w_2. Since v_1 and v_2 are orthogonal, they are linearly independent. It follows that

$$\text{span}\{v_1, v_2\} = \text{span}\{w_1, w_2\}.$$

In summary, computing v_1 and v_2 using (10.2.6), (10.2.7), and (10.2.8) yields an orthonormal basis for the plane spanned by w_1 and w_2.

The General Case

Theorem 10.2.3 *(Gram-Schmidt Orthonormalization) Let $w_1, \ldots, w_k$ be a basis for the subspace $W \subset \mathbf{R}^n$. Define v_1 as in (10.2.6) and then define inductively*

$$v'_{j+1} = w_{j+1} - (w_{j+1} \cdot v_1)v_1 - \cdots - (w_{j+1} \cdot v_j)v_j \qquad \textbf{(10.2.9)}$$

$$v_{j+1} = \frac{1}{||v'_{j+1}||} v'_{j+1}. \qquad \textbf{(10.2.10)}$$

Then $v_1, \ldots, v_k$ is an orthonormal basis of W such that for each j,

$$span\{v_1, \ldots, v_j\} = span\{w_1, \ldots, w_j\}.$$

Proof: We assume that we have constructed orthonormal vectors $v_1, \ldots, v_j$ such that

$$\text{span}\{v_1, \ldots, v_j\} = \text{span}\{w_1, \ldots, w_j\}.$$

Our purpose is to find a unit vector v_{j+1} that is orthogonal to each v_i and that satisfies

$$\text{span}\{v_1, \ldots, v_{j+1}\} = \text{span}\{w_1, \ldots, w_{j+1}\}.$$

We construct v_{j+1} in two steps. First we find a vector v'_{j+1} that is orthogonal to each of the v_i using least squares. Let w_0 be the vector in span$\{v_1, \ldots, v_j\}$ that is nearest to w_{j+1}. Theorem 10.2.2 tells us how to make this construction. Let A be the matrix whose columns are $v_1, \ldots, v_j$. Then (10.2.4) states that the coordinates of w_0 in the v_i basis are given by $(A^tA)^{-1}A^t w_{j+1}$. But since the v_is are orthonormal, the matrix A^tA is just I_k. Hence

$$w_0 = (w_{j+1} \cdot v_1)v_1 + \cdots + (w_{j+1} \cdot v_j)v_j.$$

Second, let $v'_{j+1} = w_{j+1} - w_0$ be the vector defined in (10.2.9). We claim that $v'_{j+1} = w_{j+1} - w_0$ is orthogonal to v_k for $k \leq j$ and hence to every vector in span$\{v_1, \ldots, v_j\}$. Just calculate

$$v'_{j+1} \cdot v_k = w_{j+1} \cdot v_k - w_0 \cdot v_k = w_{j+1} \cdot v_k - w_{j+1} \cdot v_k = 0.$$

Define v_{j+1} as in (10.2.10). It follows that $v_1, \ldots, v_{j+1}$ are orthonormal and that each vector is a linear combination of $w_1, \ldots, w_{j+1}$. ◆

An Example of Orthonormalization

Let $W \subset \mathbf{R}^4$ be the subspace spanned by the vectors

$$w_1 = (1, 0, -1, 0), \quad w_2 = (2, -1, 0, 1), \quad w_3 = (0, 0, -2, 1). \qquad \textbf{(10.2.11)}$$

We find an orthonormal basis for W using Gram-Schmidt orthonormalization.

Step 1: Set

$$v_1 = \frac{1}{||w_1||}\, w_1 = \frac{1}{\sqrt{2}}(1, 0, -1, 0).$$

Step 2: Following the Gram-Schmidt process, use (10.2.9) to define

$$v_2' = w_2 - (w_2 \cdot v_1)v_1 = (2, -1, 0, 1) - \sqrt{2}\,\frac{1}{\sqrt{2}}(1, 0, -1, 0) = (1, -1, 1, 1).$$

Normalization using (10.2.10) yields

$$v_2 = \frac{1}{||v_2'||}\, v_2' = \frac{1}{2}(1, -1, 1, 1).$$

Step 3: Using (10.2.9), set

$$\begin{aligned} v_3' &= w_3 - (w_3 \cdot v_1)v_1 - (w_3 \cdot v_2)v_2 \\ &= (0, 0, -2, 1) - \sqrt{2}\,\frac{1}{\sqrt{2}}(1, 0, -1, 0) - \left(-\frac{1}{2}\right)\frac{1}{2}(1, -1, 1, 1) \\ &= \frac{1}{4}(-3, -1, -3, 5). \end{aligned}$$

Normalization using (10.2.10) yields

$$v_3 = \frac{1}{||v_3'||}\, v_3' = \frac{4}{\sqrt{44}}(-3, -1, -3, 5).$$

Hence we have constructed an orthonormal basis $\{v_1, v_2, v_3\}$ for W—namely,

$$\begin{aligned} v_1 &= \frac{1}{\sqrt{2}}(1, 0, -1, 0) &&\approx (0.7071, 0, -0.7071, 0) \\ v_2 &= \frac{1}{2}(1, -1, 1, 1) &&= (0.5, -0.5, 0.5, 0.5) \\ v_3 &= \frac{4}{\sqrt{44}}(-3, -1, -3, 5) &&\approx (-0.4523, -0.1508, -0.4523, 0.7538). \end{aligned} \tag{10.2.12}$$