Bevezetés az algebrába 2 – Ortogonalitás

Wettl Ferenc
Algebra Tanszék

BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

2016. február 22.

Skaláris szorzat másik bázisban

P Legyen $\mathcal{B} = \{(2,0,1),(1,1,0),(0,-1,1)\}$. Milyen képlettel számolható ki az $\mathbf{x}\cdot\mathbf{y}$ skaláris szorzat, ha a két vektor $\mathcal{B}$-beli koordinátás alakját ismerjük?

M Jelölje az $\mathbf{x}$ vektor $\mathcal{B}$-beli koordinátás alakját $\mathbf{x}_\mathcal{B}$, a standard alakot $\mathbf{x}_\mathcal{E}$. Ekkor $\mathbf{x}_\mathcal{E} = \mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}}\mathbf{x}_\mathcal{B}$. Így

$$\mathbf{x}\cdot\mathbf{y} = \mathbf{x}_\mathcal{E}^\mathsf{T}\mathbf{y}_\mathcal{E} = (\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}}\mathbf{x}_\mathcal{B})^\mathsf{T}\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}}\mathbf{x}_\mathcal{B} = \mathbf{x}_\mathcal{B}^\mathsf{T}(\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}}^\mathsf{T}\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}})\mathbf{x}_\mathcal{B}.$$

Esetünkben

$$\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}}^\mathsf{T}\mathbf{A}_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

K Mit tudunk az $\mathbf{A}^\mathsf{T}\mathbf{A}$ mátrixról?

Szimmetrikus, invertálható, de ez még kevés.

A skaláris szorzás alaptulajdonságai $\mathbb{R}^n$-ben

T Legyen $\mathbf{u}$, $\mathbf{v}$ és $\mathbf{w}$ az $\mathbb{R}^n$ három tetszőleges vektora, és legyen c egy teszőleges valós. Ekkor

a)	$\mathbf{u}\cdot\mathbf{v}=\mathbf{v}\cdot\mathbf{u}$	a művelet fölcserélhető (kommutatív)
b)	$\mathbf{u}\cdot(\mathbf{v}+\mathbf{w})=\mathbf{u}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{w}$	disztributív
c)	$(c\mathbf{u})\cdot\mathbf{v}=c(\mathbf{u}\cdot\mathbf{v})$	a két szorzás kompatibilis

d) $\mathbf{u}\cdot\mathbf{u}\ge 0$ és $\mathbf{u}\cdot\mathbf{u}=0$ pontosan akkor teljesül, ha $\mathbf{u}=\mathbf{0}$.

B Egyszerű, pl. az *a)*:

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1+u_2v_2+\ldots+u_nv_n\\ &= v_1u_1+v_2u_2+\ldots+v_nu_n\\ &= \mathbf{v}\cdot\mathbf{u}.\end{aligned}$$

Távolság és szög $\mathbb{R}^n$-ben

D Legyen **u** és **v** az $\mathbb{R}^n$ tér két tetszőleges vektora.

1. Az **u** vektor hosszán (abszolút értékén, normáján) önmagával vett skaláris szorzatának gyökét értjük, azaz
$$|\mathbf{u}| = \|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}}. \tag{1}$$
2. Az **u** és **v** vektorok (hajlás)szögének koszinuszát az alábbi törttel definiáljuk:
$$\cos(\mathbf{u},\mathbf{v})_\angle := \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \tag{2}$$
3. Azt mondjuk, hogy az **u** és **v** vektorok merőlegesek egymásra, ha
$$\mathbf{u}\cdot\mathbf{v} = 0. \tag{3}$$
4. A két vektor végpontjának távolságán, amit egyszerűen a két vektor távolságának nevezünk, a
$$\mathrm{d}(\mathbf{u},\mathbf{v}) = |\mathbf{u}-\mathbf{v}| \tag{4}$$
értéket értjük.

Koordinátás alakok

Á $\mathbb{R}^3$-höz hasonlóan igazolhatóak:

$$|\mathbf{u}| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2},$$

$$\cos(\mathbf{u},\mathbf{v})_\angle = \frac{u_1v_1 + u_2v_2 + \ldots + u_nv_n}{\sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}\sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}}.$$

P $\mathbf{u} = (2,3,4,14)$, $\mathbf{v} = (4,6,-10,10)$, $\mathbf{w} = (0,3,6,-2)$, $|\mathbf{u}| =?$, $\mathrm{d}(\mathbf{u},\mathbf{v}) =?$, $\cos(\mathbf{u},\mathbf{w})_\angle =?$

M Az (1), a (4) és a (2) képletekkel:

$$|\mathbf{u}| = \sqrt{2^2 + 3^2 + 4^2 + 14^2} = \sqrt{225} = 15,$$

$$\begin{aligned} d(\mathbf{u},\mathbf{v}) &= \sqrt{(2-4)^2 + (3-6)^2 + (4-(-10))^2 + (14-10)^2} \\ &= \sqrt{2^2 + 3^2 + 14^2 + 4^2} = 15 \end{aligned}$$

$$\cos(\mathbf{u},\mathbf{w})_\angle = \frac{2\cdot 0 + 3\cdot 3 + 4\cdot 6 + 14\cdot(-2)}{15\cdot\sqrt{0^2 + 3^2 + 6^2 + (-2)^2}} = \frac{1}{21}.$$

Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség

K Biztosan jó a $\cos(\mathbf{u}, \mathbf{v})_\angle$ definíciója?
Ellenőrizni kéne: $|\mathbf{u} \cdot \mathbf{v}| \le |\mathbf{u}||\mathbf{v}|$ (CBS-egyenlőtlenség).

T CBS-egyenlőtlenség Tetszőleges $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vektorokra

$$|\mathbf{u} \cdot \mathbf{v}| \le |\mathbf{u}||\mathbf{v}|. \tag{5}$$

Egyenlőség pontosan akkor áll fenn, ha $\mathbf{u}$ és $\mathbf{v}$ lineárisan összefüggők, azaz ha egyik vektor a másik skalárszorosa.

B Ha $\mathbf{v} = \mathbf{0}$, akkor ✓.
Ha $\mathbf{v} \ne \mathbf{0}$, akkor legyen $\mathbf{e} = {}^{\mathbf{v}}/_{|\mathbf{v}|}$.

$$\begin{aligned} 0 &\le |\mathbf{u} - (\mathbf{u} \cdot \mathbf{e})\mathbf{e}|^2 = |\mathbf{u}|^2 + |\mathbf{u} \cdot \mathbf{e}|^2 - 2\,|\mathbf{u} \cdot \mathbf{e}|^2 = |\mathbf{u}|^2 - |\mathbf{u} \cdot \mathbf{e}|^2 \\ &= |\mathbf{u}|^2 - \frac{|\mathbf{u} \cdot \mathbf{v}|^2}{|\mathbf{v}|^2}, \end{aligned}$$

$0 = |\mathbf{u} - (\mathbf{u} \cdot \mathbf{e})\mathbf{e}|$ csak akkor állhat fönn, ha $\mathbf{u} = (\mathbf{u} \cdot \mathbf{e})\mathbf{e}$, azaz ha $\mathbf{u}$ és $\mathbf{e}$ párhuzamosak, azaz ha $\mathbf{u}$ a $\mathbf{v}$ skalárszorosa.

Háromszög-egyenlőtlenség $\mathbb{R}^n$-ben

T Tetszőleges $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vektorokra

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

B HF

Skaláris szorzat és abszolút érték (norma) kapcsolata

T Tetszőleges $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vektorokra

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2\right) \tag{6}$$

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}|^2 - |\mathbf{v}|^2\right) \tag{7}$$

B Az abszolút érték (1)-beli definíciója alapján

$$\begin{aligned}\frac{1}{4}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2\right) &= \frac{1}{4}((\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) - (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}))\\ &= \frac{1}{4}(\mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}-\mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}-\mathbf{v}\cdot\mathbf{v})\\ &= \frac{1}{4}(4\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}.\end{aligned}$$

A másik formula hasonlóan bizonyítható.

OR és ONR lineáris függetlensége

D A páronként merőleges vektorokból álló vektorrendszert ortogonális rendszernek (OR), az egységvektorokból álló OR-t ortonormált rendszernek (ONR) nevezzük.

Á Tekintsük a $\mathbf{v}_1$, $\mathbf{v}_2, \ldots,$ $\mathbf{v}_k$ vektorok rendszerét! Ha e vektorok

1. zérusvektortól különbözőek és OR-t alkotnak, vagy ha
2. ONR-t alkotnak,

akkor függetlenek.

B TFH valamely c_1, $c_2, \ldots,$ c_k konstansokra

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Mivel $i \neq j$ esetén $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, ezért a $\mathbf{v}_i$ vektorral beszorozva

$$c_i\mathbf{v}_i \cdot \mathbf{v}_i = 0,$$

amiből $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$ miatt következik, hogy $c_i = 0$.

K Egy zérusvektort nem tartalmazó OR, vagy ONR mindig bázisa az általa kifeszített altérnek.

Legjobb közelítés ONB esetén

T Adva van az $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\} \subset \mathbb{R}^n$ ONR és a $\mathbf{v} \in \mathbb{R}^n$ vektor. Ekkor a

$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}_2)\mathbf{e}_2 + \dots + (\mathbf{v} \cdot \mathbf{e}_k)\mathbf{e}_k \tag{8}$$

vektor az $\mathcal{A}$ altér $\mathbf{v}$-hez legközelebb fekvő pontja, azaz $\hat{\mathbf{v}} = \mathbf{proj}_{\mathcal{A}}\,\mathbf{v}$.

B $(\mathbf{v} - \hat{\mathbf{v}})^2 = \left(\mathbf{v} - \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)\mathbf{e}_i\right)^2 = \mathbf{v}^2 - 2\sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)^2 + \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)^2 =$ $\mathbf{v}^2 - \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)^2$. Tetszőleges $\mathbf{u}$ vektorra e távolságnégyzet:
$(\mathbf{v} - \mathbf{u})^2 = \left(\mathbf{v} - \sum_{i=1}^{k} c_i\mathbf{e}_i\right)^2 = \mathbf{v}^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^{k} c_i^2$. Különbségük
$(\mathbf{v}-\mathbf{u})^2 - (\mathbf{v}-\hat{\mathbf{v}})^2 = \left(\mathbf{v}^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v}\cdot\mathbf{e}_i) + \sum_{i=1}^{k} c_i^2\right) - \left(\mathbf{v}^2 - \sum_{i=1}^{k} (\mathbf{v}\cdot\mathbf{e}_i)^2\right) =$
$\sum_{i=1}^{k} c_i^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)^2 = \sum_{i=1}^{k} (c_i - \mathbf{v} \cdot \mathbf{e}_i)^2 \ge 0.$
Tehát a legjobb közelítés tétele szerint $\hat{\mathbf{v}} = \mathbf{proj}_{\mathcal{A}}\,\mathbf{v}$.

Legjobb közelítés ONB esetén 2.

m A $\mathbf{v}\cdot\mathbf{e}_i$ skalárt a $\mathbf{v}$ vektor $\mathbf{e}_i$-hez tartozó Fourier-együtthatójának is nevezik.

P Határozzuk meg a $(3,1,2)$ pontnak az $(2,3,6)$ és $(3,-6,2)$ vektorok által kifeszített síkra való merőleges vetületét!

K Hányféle módon tudnánk megoldani korábbi tudásunkra építve?

M E két vektor a síkban OR-t alkot! Normálás után $\mathbf{a}=\frac{1}{7}(2,3,6)$ és $\mathbf{b}=\frac{1}{7}(3,-6,2)$ ONR.

Behelyettesítés:

$$\begin{aligned}
\hat{\mathbf{v}} &= (\mathbf{v}\cdot\mathbf{a})\mathbf{a}+(\mathbf{v}\cdot\mathbf{b})\mathbf{b}\\
&= \Big((3,1,2)\cdot(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7})\Big)(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7}))+\Big((3,1,2)\cdot(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7})\Big)(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7})\\
&= 3(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7})+1(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7})\\
&= (\tfrac{9}{7},\tfrac{3}{7},\tfrac{20}{7}).
\end{aligned}$$

Ortogonális és szemiortogonális mátrixok

D Egy valós négyzetes mátrixot **ortogonálisnak** nevezünk, ha oszlopvektorai vagy sorvektorai ONR-t alkotnak. Ha nem kötjük ki, hogy a mátrix négyzetes legyen, **szemiortogonális** mátrixról beszélünk.

m Szerencsétlen szóhasználat (ortogonális – ortonormált).

m Minden ortogonális mátrix szemiortogonális is.

m Egy nem négyzetes mátrixnál vagy csak a sorai, vagy csak az oszlopai alkothatnak ONR-t, négyzetesnél mindkettő (bizonyítjuk).

m Az egységmátrix és minden permutációmátrix ortogonális.

P Melyek ortogonálisak és melyek szemiortogonálisak?

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ \mathbf{C} = \frac{1}{7}\begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}.$$

M Mindhárom mátrix szemiortogonális, **C** ortogonális.

Szemiortogonális mátrixok

T Legyen $m \geq n$ és $\mathbf{Q} \in \mathbb{R}^{m\times n}$. Az alábbi állítások ekvivalensek:

1. $\mathbf{Q}$ szemiortogonális,
2. $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_n$.

m Ha $m \leq n$, $\mathbf{Q}$ pontosan akkor szemiortogonális, ha $\mathbf{Q}\mathbf{Q}^\mathsf{T} = \mathbf{I}_m$.

m A 2 állítás algebrai nyelven azt mondja, hogy $m \geq n$ esetén $\mathbf{Q}$ pontosan akkor szemiortogonális, ha transzponáltja a bal oldali inverze.

B 1 $\Rightarrow$ 2: Ha $\mathbf{Q}$ szemiortogonális és $m \geq n$, akkor $\mathbf{Q}$ oszlopai alkotnak ONR-t. Legyen $\mathbf{Q} = [\mathbf{q}_1\ \mathbf{q}_2\ \dots\ \mathbf{q}_n]$. Ekkor $[\mathbf{Q}^\mathsf{T}\mathbf{Q}]_{ij} = \mathbf{q}_i^\mathsf{T}\mathbf{q}_j = \mathbf{q}_i \cdot \mathbf{q}_j$, de mivel a $\{\mathbf{q}_i\}$ vektorrendszer ortonormált, ezért $\mathbf{q}_i^2 = 1$ és $\mathbf{q}_i \cdot \mathbf{q}_j = 0$, ha $i \neq j$. Eszerint $[\mathbf{Q}^\mathsf{T}\mathbf{Q}]_{ii} = 1$, és $[\mathbf{Q}^\mathsf{T}\mathbf{Q}]_{ij} = 0$, ha $i \neq j$ és $i, j \leq k$, vagyis $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_k$.

2 $\Rightarrow$ 1: A $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}_k$ összefüggésbeli mátrixszorzást sorvektorszor oszlopvektorként tekintve épp azt kapjuk, hogy $\mathbf{q}_i^2 = 1$ és $\mathbf{q}_i \cdot \mathbf{q}_j = 0$, ha $i \neq j$, azaz a $\{\mathbf{q}_i\}$ vektorrendszer ortonormált.

Ortogonális mátrixok

T Legyen $\mathbf{Q} \in \mathbb{R}^{n\times n}$. Az alábbi állítások ekvivalensek:

1. $\mathbf{Q}$ oszlopvektorai ortonormált rendszert alkotnak.
2. $\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{I}_n$.
3. $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}$.
4. $\mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \mathbf{I}_n$.
5. $\mathbf{Q}$ sorvektorai ortonormált rendszert alkotnak.

B 1 $\Rightarrow$ 2: az előző állításban láttuk.

2 $\Rightarrow$ 3: $\mathbf{Q}$ négyzetes $\rightsquigarrow$ $\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{I}$ miatt $\mathbf{Q}$ invertálható $\rightsquigarrow$ $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}$.

3 $\Rightarrow$ 4: $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}} \rightsquigarrow \mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \mathbf{I}_n$.

4 $\Rightarrow$ 5: $\mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \mathbf{I}_n$ (sorvektorszor-oszlopvektor) $\rightsquigarrow$ $\mathbf{Q}$ soravektorai ONB-t alkotnak.

5 $\Rightarrow$ 1: Bizonyítottuk, hogy 1 $\Rightarrow$ 5. Alkalmazzuk ezt $\mathbf{Q}^{\mathsf{T}}$-ra.

Ortogonális mátrix inverze a transzponáltja

- Számítsuk ki az

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}, \quad \frac{1}{7}\begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}.$$

mátrixok inverzét!

M Mindhárom mátrix ortogonális, tehát az inverzek:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \quad \frac{1}{7}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}.$$

Ortogonális mátrixhoz tartozó mátrixleképezés

T Legyen $\mathbf{Q} \in \mathbb{R}^{n\times n}$. Az alábbi állítások ekvivalensek:

1. $\mathbf{Q}$ ortogonális.
2. $|\mathbf{Qx}| = |\mathbf{x}|$ minden $\mathbf{x} \in \mathbb{R}^n$ vektorra.
3. $\mathbf{Qx} \cdot \mathbf{Qy} = \mathbf{x} \cdot \mathbf{y}$ minden $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorra.

B 1 $\Rightarrow$ 2: Ha $\mathbf{Q}$ ortogonális, akkor minden $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorra

$$|\mathbf{Qx}|^2 = \mathbf{Qx} \cdot \mathbf{Qx} = (\mathbf{Qx})^\mathsf{T}(\mathbf{Qx}) = \mathbf{x}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{Qx} = \mathbf{x}^\mathsf{T}\mathbf{x} = |\mathbf{x}|^2.$$

2 $\Rightarrow$ 3: Mivel $|\mathbf{Q}(\mathbf{x}+\mathbf{y})| = |\mathbf{x}+\mathbf{y}|$ és $|\mathbf{Q}(\mathbf{x}-\mathbf{y})| = |\mathbf{x}-\mathbf{y}|$, ezért minden $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorra

$$\mathbf{Qx} \cdot \mathbf{Qy} = \frac{1}{4}\left(|\mathbf{Qx}+\mathbf{Qy}|^2 - |\mathbf{Qx}-\mathbf{Qy}|^2\right) = \frac{1}{4}\left(|\mathbf{Q}(\mathbf{x}+\mathbf{y})|^2 - |\mathbf{Q}(\mathbf{x}-\mathbf{y})\right.$$
$$= \frac{1}{4}\left(|\mathbf{x}+\mathbf{y}|^2 - |\mathbf{x}-\mathbf{y}|^2\right) = \mathbf{x} \cdot \mathbf{y}$$

3 $\Rightarrow$ 1: $\mathbf{q}_i = \mathbf{Qe}_i$ jelöléssel $\mathbf{q}_i \cdot \mathbf{q}_j = \mathbf{Qe}_i \cdot \mathbf{Qe}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

Ortogonális mátrix további tulajdonságai

T

1. Ha $\mathbf{Q}$ valós ortogonális mátrix, akkor $|\det(\mathbf{Q})| = 1$.
2. Az $n \times n$-es valós ortogonális mátrixok $O(n)$ halmazából nem vezet ki a mátrixszorzás és invertálás művelete.

B

1. $\det(\mathbf{Q}^\mathsf{T})\det(\mathbf{Q}) = \det(\mathbf{Q}^\mathsf{T}\mathbf{Q}) = \det(\mathbf{I}) = 1$, $\det(\mathbf{Q}^\mathsf{T}) = \det(\mathbf{Q}) \rightsquigarrow$ $\det(\mathbf{Q}) = 1$ vagy $\det(\mathbf{Q}) = -1$.
2. Ortogonális mátrix inverze megegyezik transzponáltjával, ami ugyancsak ortogonális, tehát inverze is az. Ha $\mathbf{Q}_1$ és $\mathbf{Q}_2$ ortogonális, akkor
$$(\mathbf{Q}_1\mathbf{Q}_2)^\mathsf{T}\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2^\mathsf{T}\mathbf{Q}_1^\mathsf{T}\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2^\mathsf{T}\mathbf{Q}_2 = \mathbf{I},$$
tehát $\mathbf{Q}_1\mathbf{Q}_2$ ortogonális.

m A második állítás azt jelenti, hogy $O(n)$ elemei csoportot alkotnak a mátrixszorzásra nézve.

m Az $n \times n$-es 1 determinánsú valós ortogonális mátrixok halmazából sem vezet ki a mátrixszorzás és invertálás művelete, tehát e mátrixok is csoportot alkotnak, e csoport jele $SO(n)$.

A sík ortogonális transzformációi

T Minden $O(2)$-be eső ortogonális mátrix vagy egy forgatás, vagy egy egyenesre való tükrözés mátrixa.

B Ha $\mathbf{Q} = \left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$ ortogonális, akkor

$$\begin{aligned} a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ ac + bd &= 0. \end{aligned} \tag{9}$$

(9) $\rightsquigarrow a^2c^2 = b^2d^2 \rightsquigarrow a^2(1-d^2) = (1-a^2)d^2 \rightsquigarrow a^2 = d^2,\ b^2 = c^2$
$\rightsquigarrow d = a$ és $c = -b$ (ekkor $\det(\mathbf{Q}) = ad - bc = 1$), vagy
$d = -a$ és $c = b$ (ekkor $\det(\mathbf{Q}) = -1$).
$\alpha \in [0, 2\pi)$ egyértelmű, melyre $a = \cos\alpha$ és $b = \sin\alpha$.

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \quad \text{vagy} \quad \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$$

egy α szögű forgatás, és egy $\alpha/2$ szögű egyenesre való tükrözés mátrixa.

m $SO(3)$ minden eleme forgatás mátrixa, $O(3) - SO(3)$ eleme egy origóra való tükrözés és egy forgatás szorzata.

P Az 1 determinánsú ortogonális

$$\frac{1}{15}\begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}$$

mátrix milyen tengely körüli és mekkora szöggel való forgatás mátrixa?

M A forgástengely irányvektorára $\mathbf{Ax} = \mathbf{x}$, azaz $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$.
Innen a forgástengely:

$$\frac{1}{15}\begin{bmatrix} -1 & -5 & 2 \\ 5 & -5 & -10 \\ 2 & 10 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Forgásszög: egy $\mathbf{v}$-re merőleges vektort képével (pl. $\mathbf{w} = (0, 1, 0)$):

$$\frac{1}{15}\begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix} \rightsquigarrow \cos\alpha = \frac{\mathbf{w}\cdot\mathbf{Aw}}{|\mathbf{w}||\mathbf{Aw}|} = \frac{2}{3}.$$

Givens-forgatás

m Az n-dimenziós tér forgatásai és tükrözései közül kiválaszthatunk olyan egyszerű, ún. primitív ortogonális transzformációkat, melyek mátrixai szorzataként az összes ortogonális mátrix előállítható.

D Givens-forgatás: forgatás, mely egy koordinátasík vektorain kívül minden más vektort helyben hagy:

$$\mathbf{G} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & \cos\alpha & \dots & -\sin\alpha & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sin\alpha & \dots & \cos\alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Givens-forgatás

P Keressük meg azt a forgatást, mely az (a, b) vektort az $(r, 0)$-ba viszi, ahol $r^2 = a^2 + b^2$.

M

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \cos\alpha &= {}^{a}\!/_{r} \\ \sin\alpha &= -{}^{b}\!/_{r} \end{aligned} \tag{10}$$

m Egy ilyen részmátrixot tartalmazó **G** Givens-forgatással elérhető, hogy egy **X** mátrix egy eleme helyén 0 legyen a **GX**-ben. (ritka mátrixok, párhuzamosítható számítások)

Householder-tükrözés

D Householder-tükrözés: hipersíkra való tükrözés. Mátrixa ($\mathbf{a}$ normálvekorral)

$$\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{a}^\mathsf{T}\mathbf{a}}\mathbf{a}\mathbf{a}^\mathsf{T}$$

Á $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} \neq \mathbf{b}$, $|\mathbf{a}| = |\mathbf{b}|$, akkor az $(\mathbf{a} - \mathbf{b})^\perp$ hipersíkra való Householder-tükrözés az $\mathbf{a}$ vektort $\mathbf{b}$-be viszi és viszont.

B A tükrözés mátrixa: $\mathbf{H} = \mathbf{I} - \frac{2}{(\mathbf{a}-\mathbf{b})^\mathsf{T}(\mathbf{a}-\mathbf{b})}(\mathbf{a} - \mathbf{b})(\mathbf{a} - \mathbf{b})^\mathsf{T}$.

Mivel $|\mathbf{a}| = |\mathbf{b}|$, ezért $\mathbf{a}^\mathsf{T}\mathbf{a} = \mathbf{b}^\mathsf{T}\mathbf{b}$, továbbá $\mathbf{a}^\mathsf{T}\mathbf{b} = \mathbf{b}^\mathsf{T}\mathbf{a}$, így $(\mathbf{a}-\mathbf{b})^\mathsf{T}(\mathbf{a}-\mathbf{b}) = \mathbf{a}^\mathsf{T}\mathbf{a}-\mathbf{a}^\mathsf{T}\mathbf{b}-\mathbf{b}^\mathsf{T}\mathbf{a}+\mathbf{b}^\mathsf{T}\mathbf{b} = 2(\mathbf{a}^\mathsf{T}\mathbf{a}-\mathbf{b}^\mathsf{T}\mathbf{a}) = 2(\mathbf{a}-\mathbf{b})^\mathsf{T}\mathbf{a}$. Eszerint

$$\begin{aligned}\mathbf{H}\mathbf{a} &= \mathbf{a} - \frac{2}{(\mathbf{a}-\mathbf{b})^\mathsf{T}(\mathbf{a}-\mathbf{b})}(\mathbf{a} - \mathbf{b})(\mathbf{a} - \mathbf{b})^\mathsf{T}\mathbf{a} \\ &= \mathbf{a} - \frac{1}{(\mathbf{a}-\mathbf{b})^\mathsf{T}\mathbf{a}}(\mathbf{a} - \mathbf{b})^\mathsf{T}\mathbf{a}(\mathbf{a} - \mathbf{b}) = \mathbf{a} - (\mathbf{a} - \mathbf{b}) = \mathbf{b}.\end{aligned}$$

Mivel $\mathbf{H}^{-1} = \mathbf{H}$, ezért $\mathbf{H}\mathbf{b} = \mathbf{H}^{-1}\mathbf{b} = \mathbf{a}$.

Householder-tükrözés

P Határozzuk meg azt a $\mathbf{H}$ mátrixot, mely az $(1,-1,-1,1)$ vektort a $(2,0,0,0)$-ba viszi.

M Az $(1,-1,-1,1)-(2,0,0,0)=(-1,-1,-1,1)$ vektorra merőleges hipersíkra való tükrözés mátrixa

$$\mathbf{H}=\mathbf{I}-\frac{2}{\mathbf{a}^\mathsf{T}\mathbf{a}}\mathbf{a}\mathbf{a}^\mathsf{T}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}-1\\-1\\-1\\1\end{bmatrix}\begin{bmatrix}-1&-1&-1&1\end{bmatrix}$$

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

Valóban, $\mathbf{H}\cdot(1,-1,-1,1)=(2,0,0,0)$.

Gram–Schmidt-ortogonalizáció

T Ha $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ egy független vektorrendszer, akkor létezik olyan ortogonális $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ vektorrendszer, hogy minden $i = 1, 2, \dots, k$ esetén

$$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i). \tag{11}$$

Ebből normálással ortonormált rendszert kapunk:

$$\left\{ \frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_k}{|\mathbf{v}_k|} \right\}$$

m Igazolható, hogy a Gram–Schmidt-ortogonalizáció működik összefüggő vektorokból álló vektorrendszerre is, annyi változással, hogy pontosan akkor lesz $\mathbf{v}_i = \mathbf{0}$, ha $\mathbf{a}_i$ nem független a kisebb indexű vektoroktól, azaz $\mathbf{a}_i$ benne van a $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i-1})$ altérben.

B A $\operatorname{span}(\mathbf{a}_1) = \operatorname{span}(\mathbf{v}_1)$ összefüggés teljesül, ha $\mathbf{v}_1 = \mathbf{a}_1$.

$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ és $\mathbf{v}_1 \perp \mathbf{v}_2$ teljesül, ha $\mathbf{v}_2$ az $\mathbf{a}_2$-nek a $\mathbf{v}_1$ által kifeszített altérre merőleges összetevője:

$$\mathbf{v}_2 = \mathbf{a}_2 - \left(\mathbf{a}_2 \cdot \frac{\mathbf{v}_1}{|\mathbf{v}_1|}\right) \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$\mathbf{v}_2 \neq \mathbf{0}$, hisz $\mathbf{v}_2 = 0$ esetén $\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{a}_1$ lenne, azaz $\mathbf{a}_1$ és $\mathbf{a}_2$ nem lenne független, ez ellentmondás.

$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ ✓

$\mathbf{v}_i \Rightarrow \mathbf{v}_{i+1}$: az $\mathbf{a}_{i+1}$ vektornak a $\operatorname{span}\left(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_i}{|\mathbf{v}_i|}\right)$ altérre merőleges összetevője legyen $\mathbf{v}_{i+1}$

$$\mathbf{v}_{i+1} = \mathbf{a}_{i+1} - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \dots - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}\mathbf{v}_i$$

$\mathbf{v}_{i+1} \neq \mathbf{0}$, különben $\mathcal{A}$ nem volna független.

$\mathbf{v}_{i+1} \in \operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i+1})$, $\mathbf{a}_{i+1} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i+1}) \rightsquigarrow (11)$ ✓

Gram–Schmidt-ortogonalizáció: 1. példa

P Ortogonalizáljuk a $\{(3,6,2),(1,9,-4),(1,2,3)\}$ vektorrendszert! Adjuk meg a tér ortonormált bázisát is!

M $\mathbf{v}_1 = (3,6,2)$

$$\mathbf{v}_2 = (1,9,-4) - \frac{(1,9,-4)\cdot(3,6,2)}{(3,6,2)\cdot(3,6,2)}(3,6,2) = (-2,3,-6)$$

$$\mathbf{v}_3 = (1,2,3) - \frac{(1,2,3)\cdot(3,6,2)}{(3,6,2)\cdot(3,6,2)}(3,6,2) - \frac{(1,2,3)\cdot(-2,3,-6)}{(-2,3,-6)\cdot(-2,3,-6)}(-2,3,-6) = \frac{1}{7}(-6,2,3)$$

Az ONR:

$$\left\{\frac{1}{7}(3,6,2), \frac{1}{7}(-2,3,-6), \frac{1}{7}(-6,2,3)\right\}$$

Gram–Schmidt-ortogonalizáció: 2. példa

P Keressünk ortonormált bázist az $(1,1,1,1)$, $(3,-1,3,-1)$, $(6,2,2,-2)$ vektorok által kifeszített altérben.

M OR:

$$\mathbf{v}_1 = (1,1,1,1)$$

$$\mathbf{v}_2 = (3,-1,3,-1) - \frac{(3,-1,3,-1)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) = (2,-2,2,-2)$$

$$\mathbf{v}_3 = (6,2,2,-2) - \frac{(6,2,2,-2)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) - \frac{(6,2,2,-2)\cdot(2,-2,2,-2)}{(2,-2,2,-2)\cdot(2,-2,2,-2)}(2,-2,2,-2) = (2,2,-2,-2)$$

Az ONR (ONB):

$$\left\{\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)\right\}$$

Gram–Schmidt-ortogonalizáció: 3. példa

P Keressünk ortonormált bázist az $(1,1,1,1)$, $(3,-1,3,-1)$, $(4,0,4,0)$, $(6,2,2,-2)$ vektorok által kifeszített altérben.

M OR:

$$\mathbf{v}_1 = (1,1,1,1)$$

$$\mathbf{v}_2 = (3,-1,3,-1) - \frac{(3,-1,3,-1)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) = (2,-2,2,-2)$$

$$\mathbf{v}_3 = (4,0,4,0) - \frac{(4,0,4,0)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) - \frac{(4,0,4,0)\cdot(2,-2,2,-2)}{(2,-2,2,-2)\cdot(2,-2,2,-2)}(2,-2,2,-2) = (0,0,0,0)$$

$$\mathbf{v}_3 = (6,2,2,-2) - \frac{(6,2,2,-2)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) - \frac{(6,2,2,-2)\cdot(2,-2,2,-2)}{(2,-2,2,-2)\cdot(2,-2,2,-2)}(2,-2,2,-2) = (2,2,-2,-2)$$

ONR, mint az előző példában.

Néhány következmény

T Bármely $\mathcal{W} \leqslant \mathbb{R}^n$ altér ONB-a kiegészíthető $\mathbb{R}^n$ ONB-ává.

B Kiegészítjük a $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\} \subset \mathcal{W}$ ONB-t az $\mathbb{R}^n$ bázisává:
$\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$,
majd azon elvégezzük a Gram–Schmidt-eljárást: ez $\mathcal{W}$ ONB-át helyben hagyja: $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$

K L! $\mathcal{W} \leqslant \mathbb{R}^n$:

1. $\mathcal{W} \oplus \mathcal{W}^\perp = \mathbb{R}^n$,
2. $(\mathcal{W}^\perp)^\perp = \mathcal{W}$.

B
1. L! $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ a fent megkonstruált bázis, így $\mathcal{W} = \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$.

   Á $\mathcal{W}^\perp = \operatorname{span}(\mathbf{e}_{k+1}, \dots, \mathbf{e}_n)$
   $\mathcal{W}^\perp \geqslant \operatorname{span}(\mathbf{e}_{k+1}, \dots, \mathbf{e}_n)$ nyilvánvaló $\rightsquigarrow \mathcal{W} + \mathcal{W}^\perp = \mathbb{R}^n$,
   másrészt $\mathcal{W} \cap \mathcal{W}^\perp = \{\mathbf{0}\} \rightsquigarrow \mathcal{W} \oplus \mathcal{W}^\perp = \mathbb{R}^n$
2. $(\mathcal{W}^\perp)^\perp \geqslant \mathcal{W}$ ✓
   $\dim(\mathcal{W}) + \dim(\mathcal{W}^\perp) = n$, $\dim(\mathcal{W}^\perp) + \dim((\mathcal{W}^\perp)^\perp) = n \rightsquigarrow$
   $\dim(\mathcal{W}) = \dim((\mathcal{W}^\perp)^\perp) \rightsquigarrow (\mathcal{W}^\perp)^\perp = \mathcal{W}$.

Legjobb közelítés ONB esetén 2

K Ha $\mathcal{W} \leqslant \mathbb{R}^n$ egy ONB-a $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\}$, akkor a

$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}_2)\mathbf{e}_2 + \cdots + (\mathbf{v} \cdot \mathbf{e}_k)\mathbf{e}_k$$

vektor az $\mathcal{A}$ altér $\mathbf{v}$-hez legközelebb fekvő pontja, azaz $\hat{\mathbf{v}} = \mathbf{proj}_{\mathcal{A}}\, \mathbf{v}$.

B (Új biz.)

$$\left(\mathbf{v} - \sum_{i=1}^{k} (\mathbf{v} \cdot \mathbf{e}_i)\mathbf{e}_i\right) \cdot \mathbf{e}_j = 0 \quad (j = 1, 2, \dots, k),$$

tehát $\mathbf{v} - \hat{\mathbf{v}} \perp \mathcal{W}$, azaz $\mathbf{v} - \hat{\mathbf{v}} \in \mathcal{W}^\perp$.

Ez a merőleges vetület definíciója szerint épp azt jelenti, hogy $\hat{\mathbf{v}} = \mathbf{proj}_{\mathcal{A}}\, \mathbf{v}$.

QR-felbontás

m Ahogyan egy mátrix elemi sorműveletekkel való háromszögalakra hozását tömör formában őrzi az LU-felbontás, ugyanígy a QR-felbontás őrzi az ortogonalizációs eljárás eredményét.

D Legyen $\mathbf{A}$ egy teljes oszloprangú mátrix. Az $\mathbf{A} = \mathbf{QR}$ felbontást QR-felbontásnak vagy redukált QR-felbontásnak nevezzük, ha $\mathbf{Q}$ az $\mathbf{A}$-val azonos méretű szemiortogonális mátrix, és $\mathbf{R}$ négyzetes felső háromszögmátrix, főátlójában pozitív elemekkel.

m A $\mathbf{Q}$ mátrixot új oszlopvektorok hozzávételével kiegészítjük egy ortogonális mátrixszá, az $\mathbf{R}$ mátrixot zérussorok hozzávételével egy $m \times n$-es felső háromszögmátrixszá, akkor ezek szorzata is $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \hat{\mathbf{Q}} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{O} \end{bmatrix} = \mathbf{QR} + \hat{\mathbf{Q}}\mathbf{O} = \mathbf{QR}.$$

Ez a teljes QR-felbontás (más ezt nevezi QR-felbontásnak). Ekkor $\mathbf{A}$-t egy ortogonális és egy $\mathbf{A}$-val azonos méretű felső háromszögmátrix szorzatára bontjuk.

A QR-felbontás létezése és egyértelműsége

T Bármely valós, teljes oszloprangú $\mathbf{A}$ mátrixnak létezik QR-felbontása, azaz létezik egy szemiortogonális $\mathbf{Q}$ mátrix és egy $\mathbf{R}$ felső háromszögmátrix pozitív főátlóbeli elemekkel, hogy $\mathbf{A} = \mathbf{QR}$. Az így kapott felbontás egyértelmű.

B L! $\mathbf{A} = [\mathbf{a}_1 \; \mathbf{a}_2 \; \dots \; \mathbf{a}_k] \in \mathbb{R}^{n\times k}$ ($\mathbf{A}$ teljes oszloprangú $\rightsquigarrow k \le n$).

Létezés: Az ortogonalizáció egységvektorait jelölje $\mathbf{q}_i$ $(i = 1, 2, \dots, k)$, tehát $\operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_i) = \operatorname{span}(\mathbf{q}_1, \dots, \mathbf{q}_i)$. Így léteznek olyan r_{ij} skalárok, hogy

$$
\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_k &= r_{1k}\mathbf{q}_1 + r_{2k}\mathbf{q}_2 + \dots + r_{kk}\mathbf{q}_k.
\end{aligned}
\tag{12}
$$

Ezt mátrixszorzat-alakba írva

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_k] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \dots \quad \mathbf{q}_k] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{kk} \end{bmatrix} = \mathbf{QR}.$$

A Gram–Schmidt-eljárásból az is látható, hogy $r_{ii} = |\mathbf{v}_i|$, tehát $r_{ii} > 0$.
R kiszámolása: $\mathbf{A} = \mathbf{QR} \rightsquigarrow \mathbf{Q}^\mathsf{T}\mathbf{A} = \mathbf{Q}^\mathsf{T}\mathbf{QR} = \mathbf{I}_k\mathbf{R} = \mathbf{R}$, tehát

$$\mathbf{R} = \mathbf{Q}^\mathsf{T}\mathbf{A}.$$

Egyértelműség: Tfh $\mathbf{A} = \mathbf{QR} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$, ahol $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \hat{\mathbf{Q}}^\mathsf{T}\hat{\mathbf{Q}} = \mathbf{I}$.

$$\left.\begin{aligned} \mathbf{A}^\mathsf{T}\mathbf{A} &= (\mathbf{QR})^\mathsf{T}\mathbf{QR} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{QR} = \mathbf{R}^\mathsf{T}\mathbf{R}, \\ \mathbf{A}^\mathsf{T}\mathbf{A} &= (\hat{\mathbf{Q}}\hat{\mathbf{R}})^\mathsf{T}\hat{\mathbf{Q}}\hat{\mathbf{R}} = \hat{\mathbf{R}}^\mathsf{T}\hat{\mathbf{Q}}^\mathsf{T}\hat{\mathbf{Q}}\hat{\mathbf{R}} = \hat{\mathbf{R}}^\mathsf{T}\hat{\mathbf{R}}. \end{aligned}\right. \rightsquigarrow \mathbf{R}^\mathsf{T}\mathbf{R} = \hat{\mathbf{R}}^\mathsf{T}\hat{\mathbf{R}} \rightsquigarrow (\hat{\mathbf{R}}^{-1})^\mathsf{T}\mathbf{R}^\mathsf{T} = \hat{\mathbf{R}}\mathbf{R}^{-1}$$

A bal oldal alsó, a jobb oldal felső háromszögmátrix $\rightsquigarrow$ mindkét szorzat diagonális. Jelölje $\mathbf{R}$ (ill. $\hat{\mathbf{R}}$) főátlója elemeit r_i (ill. $\hat{r}_i$).
$\frac{r_i}{\hat{r}_i} = \frac{\hat{r}_i}{r_i} \rightsquigarrow r_i = \hat{r}_i$ ($r_i > 0$ és $\hat{r}_i > 0$ miatt) $\rightsquigarrow (\hat{\mathbf{R}}^{-1})^\mathsf{T}\mathbf{R}^\mathsf{T} = \hat{\mathbf{R}}\mathbf{R}^{-1} = \mathbf{I}$
$\rightsquigarrow \mathbf{R} = \hat{\mathbf{R}} \rightsquigarrow \mathbf{A} = \mathbf{QR} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ miatt $\mathbf{Q} = \hat{\mathbf{Q}}$.

QR-felbontás GS-ortogonalizációból

P Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix}$$

mátrix QR-felbontását.

QR-felbontás GS-ortogonalizációból – megoldás

M A GS második példában a $(1,1,1,1)$, $(3,-1,3,-1)$, $(6,2,2,-2)$ bázisból az ONB: $\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$, $\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)$, $\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)$

$$\mathbf{Q}=\frac{1}{2}\begin{bmatrix}1&1&1\\1&-1&1\\1&1&-1\\1&-1&-1\end{bmatrix}$$

$$\mathbf{R}=\mathbf{Q}^{\mathsf{T}}\mathbf{A}=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\end{bmatrix}\begin{bmatrix}1&3&6\\1&-1&2\\1&3&2\\1&-1&-2\end{bmatrix}=\begin{bmatrix}2&2&4\\0&4&4\\0&0&4\end{bmatrix}$$

Valóban,

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

QR-felbontás Givens-forgatásokkal

P Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & 8 \\ 3 & 10 & 6 \\ 0 & 12 & 13 \end{bmatrix}$$

mátrix QR-felbontását Givens-forgatások segítségével!

M az első és második sorokat és oszlopokat figyelve elimináljuk a második sor első elemét: $a = 4$, $b = 3$, tehát $r = \sqrt{3^2 + 4^2} = 5$, $\cos\alpha = 4/5$, $\sin\alpha = -3/5$.

$$\mathbf{Q}_1 = \begin{bmatrix} 4/5 & 3/5 & 0 \\ -3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 5 & 0 \\ 0 & 12 & 13 \end{bmatrix}.$$

Következő lépésben a $\mathbf{Q}_1\mathbf{A}$ mátrix harmadik sorának második elemét elimináljuk:

$$\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5/13 & 12/13 \\ 0 & -12/13 & 5/13 \end{bmatrix}. \qquad \mathbf{R} = \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 13 & 12 \\ 0 & 0 & 5 \end{bmatrix}.$$

és innen

$$\mathbf{Q} = (\mathbf{Q}_2\mathbf{Q}_1)^{-1} = \mathbf{Q}_1^\mathsf{T}\mathbf{Q}_2^\mathsf{T} = \begin{bmatrix} 4/5 & -3/13 & 36/65 \\ 3/5 & 4/13 & -48/65 \\ 0 & 12/13 & 5/13 \end{bmatrix},$$

amely mátrixokkal $\mathbf{A} = \mathbf{QR}$ valóban fennáll.

QR-felbontás Hauseholder-tükrözésekkel

- Ötlet:

$$\mathbf{A} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \to \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \to \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

$$\mathbf{Q}_1 = \mathbf{H}_1 \qquad \mathbf{Q}_2 = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \mathbf{H}_2 & \\ 0 & & & \end{array}\right]$$

$$\to \mathbf{Q}_3\mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

$$\mathbf{Q}_3 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{H}_3$}} \\ 0 & 0 & & \end{array}\right]$$

QR-felbontás Householder-tükrözésekkel

P Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \\ -2 & 5 & -7 \end{bmatrix}$$

mátrix QR-felbontását Householder-módszerrel!

M Az $(1, 2, -2) \mapsto (3, 0, 0)$ transzformációhoz az

$$\mathbf{a} = (1, 2, -2) - (3, 0, 0) = (-2, 2, -2)$$

vektorral Householder-tükrözést végzünk:

$$\mathbf{Q}_1 = \mathbf{I}_3 - \frac{2}{\mathbf{a}^\mathsf{T}\mathbf{a}}\mathbf{a}\mathbf{a}^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 4 & -4 & 4 \\ -4 & 4 & -4 \\ 4 & -4 & 4 \end{bmatrix}$$

$$= \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$$

QR-felbontás Householder-tükrözésekkel

Ezután a $\mathbf{Q}_1\mathbf{A}$ mátrixból képzeletben elhagyva az első sort és oszlopot a $(4,3)\mapsto(5,0)$ transzformációhoz kell az $\mathbf{a}=(4,3)-(5,0)=(-1,3)$ vektorral Householder-tükrözést végezni:

$$\mathbf{H}_2=\mathbf{I}_2-\frac{2}{\mathbf{a}^\mathsf{T}\mathbf{a}}\mathbf{a}\mathbf{a}^\mathsf{T}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\frac{1}{5}\begin{bmatrix}1 & -3\\ -3 & 9\end{bmatrix}=\frac{1}{5}\begin{bmatrix}4 & 3\\ 3 & -4\end{bmatrix}$$

$$\mathbf{Q}_2=\left[\begin{array}{c|cc}1 & 0 & 0\\ \hline 0 & 4/5 & 3/5\\ 0 & 3/5 & -4/5\end{array}\right],\qquad \mathbf{R}=\mathbf{Q}_2\mathbf{Q}_1\mathbf{A}=\begin{bmatrix}3 & -2 & 3\\ 0 & 5 & -7\\ 0 & 0 & 1\end{bmatrix}$$

$$\mathbf{Q}=(\mathbf{Q}_2\mathbf{Q}_1)^{-1}=\mathbf{Q}_1^\mathsf{T}\mathbf{Q}_2^\mathsf{T}=\frac{1}{15}\begin{bmatrix}5 & 2 & 14\\ 10 & 10 & -5\\ -10 & 11 & 2\end{bmatrix}.$$

Egyenletrendszer optimális megoldása QR-felbontással

T Legyen $\mathbf{A} \in \mathbb{R}^{m\times n}$ teljes oszloprangú, $\mathbf{A} = \mathbf{QR}$ egy QR-felbontása, és $\mathbf{b} \in \mathbb{R}^n$. Ekkor az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer egyetlen optimális megoldása $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^\mathsf{T}\mathbf{b}$, ami megkapható az

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^\mathsf{T}\mathbf{b}$$

egyenletrendszerből egyszerű visszahelyettesítéssel is.

B A normálegyenletből indulva

$$\begin{aligned}
\mathbf{A}^\mathsf{T}\mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}^\mathsf{T}\mathbf{b} && \mathbf{A} = \mathbf{QR} \text{ behelyettesítése után}\\
(\mathbf{QR})^\mathsf{T}\mathbf{QR}\hat{\mathbf{x}} &= (\mathbf{QR})^\mathsf{T}\mathbf{b} && \\
\mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{QR}\hat{\mathbf{x}} &= \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} && \mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}\\
\mathbf{R}^\mathsf{T}\mathbf{R}\hat{\mathbf{x}} &= \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} && \text{balról szorzás az } (\mathbf{R}^\mathsf{T})^{-1} \text{ mátrixszal}\\
\mathbf{R}\hat{\mathbf{x}} &= \mathbf{Q}^\mathsf{T}\mathbf{b}. &&
\end{aligned}$$

Egyenletrendszer optimális megoldása QR-felbontással

P Oldjuk meg QR-felbontással:

$$\begin{aligned} x + 3y + 6z &= 8 \\ x - y + 2z &= 2 \\ x + 3y + 2z &= 2 \\ x - y - 2z &= 0 \end{aligned}$$

QR-felbontás: $\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$

$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^\mathsf{T}\mathbf{b}$: $\begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 8 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 4 \end{bmatrix}$

Megoldás visszahelyettesítéssel: $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (1, 0, 1)$.