Ortogonalisace a ortogonální projekce

Odpřednesenou látku naleznete v kapitole 12.4 skript *Abstraktní a konkrétní lineární algebra*.

Minulé přednášky

1. Definice skalárního součinu v lineárních prostorech nad $\mathbb{R}$.
2. Úplný popis skalárních součinů v prostoru $\mathbb{R}^n$.

Dnešní přednáška

1. V této přednášce (a ve všech přednáškách týkajících se skalárního součinu) se zaměříme na lineární prostory nad $\mathbb{R}$.
2. Ortogonální báze a ortonormální báze.
3. Ortogonální projekce. Ortogonalisace a ortonormalisace.

Důležitá aplikace — viz příští přednášku

1. Metoda nejmenších čtverců.

Připomenutí vlastností ortogonality (minulé přednášky)

Platí-li $\langle \vec{x} \mid \vec{y} \rangle = 0$, říkáme, že vektory $\vec{x}$ a $\vec{y}$ jsou ortogonální (také: navzájem kolmé).

1. Pozor: nulový vektor $\vec{o}$ je kolmý na každý vektor $\vec{x}$.
   Obráceně: jestliže $\vec{x}$ je kolmý na každý vektor, pak $\vec{x} = \vec{o}$.
2. Abychom ukázali, že rovnost $\langle \vec{x} \mid \vec{w} \rangle = 0$ platí pro každý vektor $\vec{w}$ ze $W = \mathrm{span}(M)$, stačí ukázat, že pro všechny vektory $\vec{m}$ z M platí rovnost $\langle \vec{x} \mid \vec{m} \rangle = 0$.

   Speciální případ výše uvedeného je:[a]

   Ať $(\vec{b}_1, \ldots, \vec{b}_k)$ je báze lineárního podprostoru W lineárního prostoru L. Jestliže platí $\langle \vec{x} \mid \vec{b}_i \rangle = 0$ pro všechna $i = 1, \ldots, k$, potom platí $\langle \vec{x} \mid \vec{w} \rangle = 0$ pro všechny vektory $\vec{w}$ z W.

[a] Tento speciální případ několikrát (bez dalších komentářů) v dnešní přednášce použijeme.

Tvrzení (lineární nezávislost ortogonální množiny vektorů)

Ať M je jakákoli množina nenulových vektorů s vlastností $\langle \vec{x} \mid \vec{y} \rangle = 0$ pro jakékoli různé vektory $\vec{x}$, $\vec{y}$ z M.[a] Pak M je lineárně nezávislá množina.

[a] Takové množině říkáme ortogonální množina.

Důkaz.

Ať $\{\vec{x}_1, \ldots, \vec{x}_n\}$ je jakákoli konečná podmnožina M. Ať $\sum_{i=1}^{n} a_i \cdot \vec{x}_i = \vec{o}$. Pro libovolné pevné $i_0 \in \{1, \ldots, n\}$ platí:

$$0 = \langle \vec{x}_{i_0} \mid \vec{o} \rangle = \langle \vec{x}_{i_0} \mid \sum_{i=1}^{n} a_i \cdot \vec{x}_i \rangle = \sum_{i=1}^{n} a_i \cdot \langle \vec{x}_{i_0} \mid \vec{x}_i \rangle = a_{i_0} \cdot \langle \vec{x}_{i_0} \mid \vec{x}_{i_0} \rangle.$$

Protože $\vec{x}_{i_0} \neq \vec{o}$, platí $\langle \vec{x}_{i_0} \mid \vec{x}_{i_0} \rangle \neq 0$. Proto $a_{i_0} = 0$. ■

Několik sloganů

1. Jestliže $\dim(L) = n$, potom každá ortogonální množina v L má nejvýše n prvků.

   Slogan: v prostoru dimense n může existovat maximálně n navzájem na sebe kolmých nenulových vektorů.

2. Ortogonální množině v L, která tvoří bázi L, říkáme ortogonální báze.

   Slogan:[a] ortogonální báze je pravoúhlý souřadnicový systém.

3. Každou bázi $(\vec{b}_1, \ldots, \vec{b}_n)$ lze normalisovat: v bázi $\left(\dfrac{\vec{b}_1}{\|\vec{b}_1\|}, \ldots, \dfrac{\vec{b}_n}{\|\vec{b}_n\|}\right)$ mají všechny vektory normu 1.

   Slogan: normální báze má jednotkové úseky na jednotlivých souřadnicových osách.

[a] Pozor: jde jen o slogan. Víme, že například leckterý skalární součin v rovině může jako ortogonální vidět vektory, které „na papíře“ nesvírají pravý úhel.

Definice (ortonormální báze, čili normální a ortogonální báze)

Bázi $(\vec{b}_1, \ldots, \vec{b}_n)$ prostoru se skalárním součinem, která splňuje rovnost $\langle \vec{b}_i \mid \vec{b}_j \rangle = \delta_{ij}$,[a] říkáme ortonormální.

[a] Kroneckerův symbol δ splňuje: $\delta_{ii} = 1$, $\delta_{ij} = 0$ pro $i \neq j$.

Poznámky

1. Kanonická báze $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ prostoru $\mathbb{R}^n$ je ortonormální vzhledem ke standardnímu skalárnímu součinu.
2. Pro jakoukoli bázi $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ prostoru $\mathbb{R}^n$ existuje jednoznačně určený skalární součin, ve kterém je tato báze ortonormální (viz minulou přednášku).

   Stejnou větu lze dokázat pro obecné prostory nad $\mathbb{R}$ konečné dimense. To dokazovat nebudeme.
3. Ortonormální báze jsou důležité: umožňují „zpříjemnit" řadu výpočtů. Viz dále.

Tvrzení (výpočet souřadnic v ortonormální bázi)

Ať $B = (\vec{b}_1, \ldots, \vec{b}_n)$ je ortonormální báze prostoru se skalárním součinem. Pak[a] $\vec{x} = \sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \vec{b}_i$, čili $\mathbf{coord}_B(\vec{x}) = \begin{pmatrix} \langle \vec{b}_1 \mid \vec{x} \rangle \\ \vdots \\ \langle \vec{b}_n \mid \vec{x} \rangle \end{pmatrix}$.

[a] V tomto kontextu se používá i terminologie: $\langle \vec{b}_i \mid \vec{x} \rangle$ je Fourierův koeficient vektoru $\vec{x}$ vzhledem k ose $\vec{b}_i$, součet $\sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \vec{b}_i$ je Fourierova řada vektoru $\vec{x}$.

Důkaz.

Označme $\vec{s} = \sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \vec{b}_i$. Musíme ukázat: $\vec{x} = \vec{s}$, tj., $\vec{x} - \vec{s} = \vec{o}$.

Pro libovolné pevné $i_0 \in \{1, \ldots, n\}$ platí $\langle \vec{b}_{i_0} \mid \vec{x} - \vec{s} \rangle = \ldots = 0$ (využíváme vlastnosti skalárního součinu a faktu, že B je ortonormální).

Takže $\vec{x} - \vec{s} = \vec{o}$, neboli $\vec{x} = \vec{s}$. ■

Důsledek (skalární součin v ortonormální bázi)

Ať $B = (\vec{b}_1, \ldots, \vec{b}_n)$ je ortonomální báze prostoru se skalárním součinem. Pak platí:[a] $\langle \vec{x} \mid \vec{y} \rangle = \sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \langle \vec{b}_i \mid \vec{y} \rangle$.

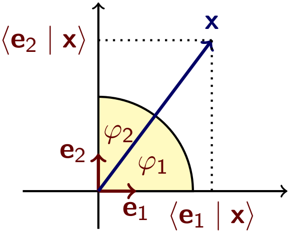
[a] Podle předchozího to znamená: $\langle \vec{x} \mid \vec{y} \rangle = \mathbf{x}^T \cdot \mathbf{y}$, kde $\mathbf{coord}_B(\vec{x}) = \mathbf{x}$ a $\mathbf{coord}_B(\vec{y}) = \mathbf{y}$. Slogan: skalární součin v ortonormální bázi se počítá jako standardní skalární součin souřadnicových vektorů. Příslušné rovnosti se říká Parsevalova rovnost.

Důkaz.

Spočteme Fourierovy řady: $\vec{x} = \sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \vec{b}_i$ a $\vec{y} = \sum_{j=1}^{n} \langle \vec{b}_j \mid \vec{y} \rangle \cdot \vec{b}_j$.

Potom (využíváme vlastnosti skalárního součinu a faktu, že B je ortonormální) $\langle \vec{x} \mid \vec{y} \rangle = \ldots = \sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle \cdot \langle \vec{b}_i \mid \vec{y} \rangle$. ∎

Direction Cosines v rovině s ortonormální bází $(\mathbf{e}_1, \mathbf{e}_2)$



$$\cos\varphi_1 = \frac{\text{orientovaná délka přilehlé odvěsny}}{\text{délka přepony}} = \frac{\langle \mathbf{e}_1 \mid \mathbf{x} \rangle}{\|\mathbf{x}\|}$$

$$\cos\varphi_2 = \frac{\text{orientovaná délka přilehlé odvěsny}}{\text{délka přepony}} = \frac{\langle \mathbf{e}_2 \mid \mathbf{x} \rangle}{\|\mathbf{x}\|}$$

$\varphi_1 + \varphi_2 = \dfrac{\pi}{2}$, čili $\cos\varphi_2 = \cos\left(\dfrac{\pi}{2} - \varphi_1\right) = \sin\varphi_1$

Tudíž $\cos^2\varphi_1 + \cos^2\varphi_2 = \cos^2\varphi_1 + \sin^2\varphi_1 = 1$.

Tvrzení (Direction Cosines)

Ať $B = (\vec{b}_1, \ldots, \vec{b}_n)$ je ortonormální báze prostoru se skalárním součinem. Ať vektor $\vec{x}$ je nenulový. Potom pro úhel φ_{i_0}, který vektor $\vec{x}$ svírá se souřadnicovou osou $\vec{b}_{i_0}$, platí rovnost[a]

$\cos\varphi_{i_0} = \dfrac{\langle \vec{b}_{i_0} \mid \vec{x} \rangle}{\|\vec{x}\|}$. Navíc platí $\displaystyle\sum_{i=1}^{n} \cos^2\varphi_i = 1$.

---

[a] Všimněme si: tvrdíme, že $\langle \vec{b}_{i_0} \mid \vec{x} \rangle$, tj. i_0-tá souřadnice vektoru $\vec{x}$ vzhledem k bázi B, se počítá jako součin $\|\vec{x}\| \cdot \cos\varphi_{i_0}$. To je zobecnění známého faktu z elementární geometrie roviny (viz předchozí stranu).

Důkaz.

Protože $\langle \vec{b}_{i_0} \mid \vec{x} \rangle = \underbrace{\|\vec{b}_{i_0}\|}_{=1} \cdot \|\vec{x}\| \cdot \cos\varphi_{i_0}$, platí $\cos\varphi_{i_0} = \dfrac{\langle \vec{b}_{i_0} \mid \vec{x} \rangle}{\|\vec{x}\|}$.

Dále: $$\sum_{i=1}^{n} \cos^2\varphi_i = \sum_{i=1}^{n} \frac{\langle \vec{b}_i \mid \vec{x} \rangle^2}{\|\vec{x}\|^2} = \frac{\displaystyle\sum_{i=1}^{n} \langle \vec{b}_i \mid \vec{x} \rangle^2}{\|\vec{x}\|^2} = \frac{\langle \vec{x} \mid \vec{x} \rangle}{\|\vec{x}\|^2} = 1. \qquad \blacksquare$$

Ortogonální projekce na přímku a ortogonální rejekce přímkou

Ať $\vec{s} \neq \vec{o}$ je vektor v prostoru L se skalárním součinem $\langle - \mid - \rangle$. Potom pro každý vektor $\vec{x}$ platí:

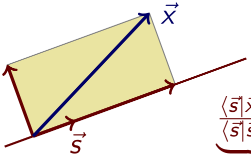
1. Vektor $\frac{\langle \vec{s} | \vec{x} \rangle}{\langle \vec{s} | \vec{s} \rangle} \cdot \vec{s}$ zjevně leží na přímce $\mathrm{span}(\vec{s})$.
2. Vektor $\vec{x} - \frac{\langle \vec{s} | \vec{x} \rangle}{\langle \vec{s} | \vec{s} \rangle} \cdot \vec{s}$ je kolmý na přímku $\vec{s}$, protože platí

$$\langle \vec{s} \mid \vec{x} - \frac{\langle \vec{s} | \vec{x} \rangle}{\langle \vec{s} | \vec{s} \rangle} \cdot \vec{s} \rangle = \langle \vec{s} \mid \vec{x} \rangle - \langle \vec{s} \mid \vec{x} \rangle \cdot \underbrace{\frac{\langle \vec{s} | \vec{s} \rangle}{\langle \vec{s} | \vec{s} \rangle}}_{=1} = 0$$

Dostáváme tedy ortogonální rozklad[a] vektoru $\vec{x}$:

$$\underbrace{\vec{x} - \frac{\langle \vec{s} | \vec{x} \rangle}{\langle \vec{s} | \vec{s} \rangle} \cdot \vec{s}}$$

ortogonální rejekce přímkou $\mathrm{span}(\vec{s})$



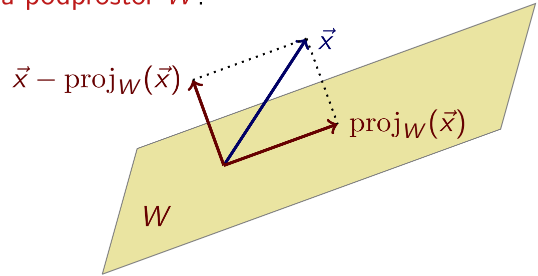
$$\underbrace{\frac{\langle \vec{s} | \vec{x} \rangle}{\langle \vec{s} | \vec{s} \rangle} \cdot \vec{s}}$$

ortogonální projekce na přímku $\mathrm{span}(\vec{s})$

[a] Slovník: projekce=promítnutí, rejekce=odmítnutí.

Zobecnění: projekce na podprostor a rejekce podprostorem

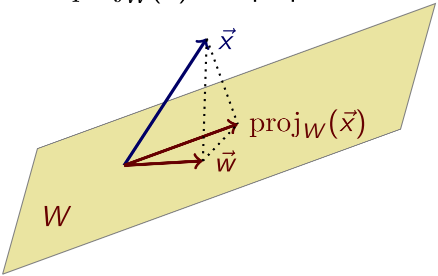
Ať W je podprostor lineárního prostoru L se skalárním součinem, ať $\vec{x}$ je libovolný vektor v L.
Vektoru $\mathrm{proj}_W(\vec{x})$, který leží ve W a pro který je $\vec{x} - \mathrm{proj}_W(\vec{x})$ kolmý na všechny vektory z W, říkáme ortogonální projekce vektoru $\vec{x}$ na podprostor W.



Vektoru $\vec{x} - \mathrm{proj}_W(\vec{x})$ budeme říkat ortogonální rejekce vektoru $\vec{x}$ podprostorem W a budeme jej značit $\mathrm{rej}_W(\vec{x})$.

Ortogonální rejekce je „nejkratší“ ze všech rejekcí

Pro jakýkoli vektor $\vec{x}$, který neleží ve W, a pro jakýkoli vektor $\vec{w}$, který ve W leží, vzniká pravoúhlý trojúhelník s odvěsnami $\mathrm{proj}_W(\vec{x}) - \vec{w}$ a $\vec{x} - \mathrm{proj}_W(\vec{x})$, a s přeponou $\vec{x} - \vec{w}$.



Díky Pythagorově větě tedy pro všechny vektory $\vec{w}$ z W platí[a]

$$\|\vec{x}-\mathrm{proj}_W(\vec{x})\|^2 \leq \|\vec{x}-\mathrm{proj}_W(\vec{x})\|^2+\underbrace{\|\mathrm{proj}_W(\vec{x}) - \vec{w}\|^2}_{\geq 0} = \|\vec{x}-\vec{w}\|^2$$

[a] Na této nerovnosti je založena metoda nejmenších čtverců, viz téma 11A.

Věta (ortogonální projekce na podprostor s ortogonální bází)
Ať $M = \{\vec{u}_1, \ldots, \vec{u}_k\}$ je konečná neprázdná ortogonální množina vektorů. Označme $W = \mathrm{span}(M)$. Pro libovolný vektor $\vec{x}$ je

$$\mathrm{proj}_W(\vec{x}) = \sum_{i=1}^{k} \mathrm{proj}_{\mathrm{span}(\vec{u}_i)}(\vec{x}) = \sum_{i=1}^{k} \frac{\langle \vec{u}_i \mid \vec{x} \rangle}{\langle \vec{u}_i \mid \vec{u}_i \rangle} \cdot \vec{u}_i$$

ortogonální projekce vektoru $\vec{x}$ na podprostor W.

Důkaz.

Označte: $\vec{p} = \displaystyle\sum_{i=1}^{k} \frac{\langle \vec{u}_i \mid \vec{x} \rangle}{\langle \vec{u}_i \mid \vec{u}_i \rangle} \cdot \vec{u}_i$. Potom platí:

1. $\vec{p}$ leží ve W.
2. Pro každé $i_0 = 1, \ldots, k$ platí $\langle \vec{u}_{i_0} \mid \vec{x} - \vec{p} \rangle = \ldots = 0$.

To znamená, že $\vec{p} = \mathrm{proj}_W(\vec{x})$. ∎

Ortogonalisační proces (Gram-Schmidt)

Každou bázi $B = (\vec{b}_1, \ldots, \vec{b}_n)$ prostoru se skalárním součinem lze převést na bázi $C = (\vec{c}_1, \ldots, \vec{c}_n)$ s následujícími vlastnostmi:

1. C je ortogonální, tj $\langle \vec{c}_i \mid \vec{c}_j \rangle = 0$ pro $i \neq j$.
2. Pro každé $k \in \{1, \ldots, n\}$ platí $\operatorname{span}\{\vec{b}_1, \ldots, \vec{b}_k\} = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_k\}$.

Důkaz.

Definujeme[a]

$$\vec{c}_1 := \vec{b}_1, \quad \vec{c}_{k+1} := \underbrace{\mathrm{rej}_{\operatorname{span}\{\vec{c}_1,\ldots,\vec{c}_k\}}(\vec{b}_{k+1})}_{\text{rejekce vektoru } \vec{b}_{k+1} \text{ podprostorem } \operatorname{span}\{\vec{c}_1,\ldots,\vec{c}_k\}}$$

Díky definici je splněno $\operatorname{span}\{\vec{b}_1, \ldots, \vec{b}_k\} = \operatorname{span}\{\vec{c}_1, \ldots, \vec{c}_k\}$, pro každé $k \in \{1, \ldots, n\}$.
První podmínka je splněna z definice ortogonální rejekce. ■

[a] Slogan: Gram-Schmidt je posloupnost postupných ortogonálních rejekcí.

Poznámka (ortonormalisační proces)

Pokud je $C = (\vec{c}_1, \ldots, \vec{c}_n)$ ortogonální báze[a] prostoru L, je seznam $(\frac{\vec{c}_1}{\|\vec{c}_1\|}, \ldots, \frac{\vec{c}_n}{\|\vec{c}_n\|})$ ortonormální báze prostoru L (tj je ortogonální a norma každého prvku je 1):

$$\langle \frac{\vec{c}_i}{\|\vec{c}_i\|} \mid \frac{\vec{c}_j}{\|\vec{c}_j\|} \rangle = \frac{1}{\|\vec{c}_i\| \cdot \|\vec{c}_j\|} \cdot \langle \vec{c}_i \mid \vec{c}_j \rangle = \delta_{ij}$$

Každou konečnou bázi B v prostoru se skalárním součinem tedy lze ortonormalisovat:

1. Nejprve provedeme Gram-Schmidtův ortogonalisační proces na bázi B. Dostaneme ortogonální bázi C.
2. Ortogonální bázi C znormalisujeme výše uvedeným postupem.

[a] Evidentně pro každé i platí $\|\vec{c}_i\| \neq 0$, protože C je báze.

Příklad (ortogonalisace vektorů — Gram-Schmidt)

Vektory $\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ a $\mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ jsou lineárně nezávislé v $\mathbb{R}^4$. Příslušnou bázi $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ podprostoru W dimense 3 označíme B.

Báze B není ortogonální vzhledem ke standardnímu skalárnímu součinu v $\mathbb{R}^4$. Bázi B nyní ortogonalisujeme. Výsledné vektory v nové (ortogonální) bázi označíme $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$. Budeme postupovat Gram-Schmidtovou metodou.

1. První vektor: $\mathbf{c}_1 = \mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

Příklad (pokrač.)

2. Druhý vektor: spočteme

$$\mathbf{b}_2 - \mathrm{proj}_{\mathsf{span}\{\mathbf{c}_1\}}(\mathbf{b}_2) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}.$$

Užitečný trik: protože skalární násobek nemění ortogonalitu,

položíme $\mathbf{c}_2 = 4 \cdot \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

Tím se zbavíme pozdějších nepříjemných výpočtů se zlomky.

Příklad (pokrač.)

3. Třetí vektor: spočteme

$$\mathbf{b}_3 - \mathrm{proj}_{\mathsf{span}\{\mathbf{c}_1,\mathbf{c}_2\}}(\mathbf{b}_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{4} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{12} \cdot \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

Opět se zbavíme zlomků: $\mathbf{c}_3 = 3 \cdot \begin{pmatrix} 0 \\ -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$.

Výpočet je u konce: seznam $(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ je hledaná ortogonální báze.

Příklad (normalisace ortogonální báze)

Normalisace ortogonální báze $\mathbf{c}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{c}_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{c}_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$ podprostoru W v prostoru $\mathbb{R}^4$ se standardním skalárním součinem je tvořena vektory

$$\frac{\mathbf{c}_1}{\|\mathbf{c}_1\|} = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{\mathbf{c}_2}{\|\mathbf{c}_2\|} = \frac{1}{\sqrt{12}} \cdot \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{\mathbf{c}_3}{\|\mathbf{c}_3\|} = \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$$