Charakterisace skalárních součinů v $\mathbb{R}^n$

Odpřednesenou látku naleznete v kapitolách 12.1, 12.2 a 12.3 skript *Abstraktní a konkrétní lineární algebra*.

Dnešní přednáška

1. V této přednášce (a ve všech přednáškách týkajících se skalárního součinu) se zaměříme na lineární prostory nad $\mathbb{R}$.
2. Charakterisace matic, které zadávají skalární součiny v prostoru $\mathbb{R}^n$.
3. Konstrukce skalárních součinů požadovaných vlastností.

Příští přednášky ke skalárnímu součinu

1. Ortogonální báze a ortonormální báze.
2. Ortogonalisační proces a ortonormalisační proces.
3. Ortogonální projekce a ortogonální rejekce.

Připomenutí (dva různé skalární součiny v $\mathbb{R}^2$)

1. Standardní skalární součin:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 + x_2 y_2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbf{E}_2} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

2. „Nezvyklý“ skalární součin:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 + x_2 y_1 + x_1 y_2 + 2 x_2 y_2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}_{\mathbf{G}} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

V obou případech je součin zadán jistou maticí typu 2×2. Je to náhoda?

Co dál?

Ukážeme, že skalární součiny v $\mathbb{R}^n$ odpovídají přesně těm čtvercovým maticím, kterým říkáme positivně definitní.

Definice (positivně definitní matice)

Řekneme, že matice $\mathbf{G}$ typu $n \times n$ nad $\mathbb{R}$ je positivně definitní, když existuje matice $\mathbf{R}$ s lineárně nezávislými sloupci tak, že $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$.

Poznámky

1. Protože $\mathbf{G}^T = (\mathbf{R}^T \cdot \mathbf{R})^T = \mathbf{R}^T \cdot \mathbf{R} = \mathbf{G}$, je každá positivně definitní matice $\mathbf{G}$ symetrická.

Poznámky (pokrač.)

2. Positivně definitní matice $\mathbf{G}$ zobecňují kladná reálná čísla: matice $\mathbf{R}$ je „druhá odmocnina“[a] matice $\mathbf{G}$.

   Opravdu: Matice $\mathbf{G} = (g)$ typu 1×1 je positivně definitní právě tehdy, když $g > 0$.

   1. Ať $\mathbf{G} = (g) = \mathbf{R}^T \cdot \mathbf{R}$. Pak $\mathbf{R} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ a platí $g = r_1^2 + \cdots + r_n^2$.

      Protože jediný sloupec $\mathbf{R}$ musí být lineárně nezávislý, je $g > 0$.
   2. Je-li $g > 0$, platí $(g) = (\sqrt{g})^T \cdot (\sqrt{g})$. Protože $\sqrt{g} > 0$, je jediný sloupec matice $\mathbf{R} = (\sqrt{g})$ lineárně nezávislý. Matice $\mathbf{G}$ je tudíž positivně definitní.

---

[a] Jde jen o slogan: matice $\mathbf{R}$ není určena jednoznačně. Například platí

$$(4) = (2)^T \cdot (2) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Věta (charakterisace positivně definitních matic)

Pro matici $\mathbf{G} = (g_{ij})_{i=1,\dots,n,j=1,\dots,n}$ nad $\mathbb{R}$ jsou následující podmínky ekvivalentní:

1. $\mathbf{G}$ je positivně definitní.
2. Matice $\mathbf{G}$ je symetrická a determinanty všech matic $\mathbf{G}_k = (g_{ij})_{i=1,\dots,k,j=1,\dots,k}$, kde $1 \leqslant k \leqslant n$, jsou kladné.[a]
3. Matice $\mathbf{G}$ je symetrická a nerovnost $\mathbf{x}^T \cdot \mathbf{G} \cdot \mathbf{x} \geqslant 0$ platí pro všechna $\mathbf{x}$ z $\mathbb{R}^n$ (rovnost platí pouze pro $\mathbf{x} = \mathbf{o}$).
4. Matice $\mathbf{G}$ je symetrická a $\mathrm{char}_{\mathbf{G}}(x)$ má všechny kořeny reálné a kladné.
5. Existuje regulární matice $\mathbf{R}$ tak, že platí $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$.

---

[a] Tento test positivní definitnosti budete využívat v analýze pro určování lokálních minim funkcí více proměnných.

Důkaz.

Bez důkazu (je těžký, pro zájemce: Tvrzení 12.3.4 skript). ■

Poznámka o Choleskyho faktorisaci — nepovinné

1. Připomenutí definice: $\mathbf{G}$ je positivně definitní právě tehdy, když $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$, kde $\mathbf{R}$ má lineárně nezávislé sloupce.
2. Předchozí věta: $\mathbf{G}$ je positivně definitní právě tehdy, když $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$, kde $\mathbf{R}$ je regulární.
3. Zesílení věty: $\mathbf{G}$ je positivně definitní právě tehdy, když $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$, kde $\mathbf{R}$ je regulární v horním blokovém tvaru.

   Rovnosti $\mathbf{G} = \mathbf{R}^T \cdot \mathbf{R}$ pro regulární matici $\mathbf{R}$ v horním blokovém tvaru se říká Choleskyho faktorisace matice $\mathbf{G}$. Příklad Choleskyho faktorisace:

$$\begin{pmatrix} 1 & 2 & 8 \\ 2 & 8 & 12 \\ 8 & 12 & 27 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

Choleskyho faktorisaci lze nalézt algoritmem, viz skripta, Příklad 12.3.6. Tento algoritmus je nepovinný.

Příklady

1. Protože $\mathbf{E}_n = \mathbf{E}_n^T \cdot \mathbf{E}_n$, je jednotková matice $\mathbf{E}_n$ positivně definitní.

   Připomeňme: $\mathbf{E}_n$ zadává standardní skalární součin $\langle \mathbf{x} \mid \mathbf{y} \rangle = \mathbf{x}^T \cdot \mathbf{E}_n \cdot \mathbf{y} = \mathbf{x}^T \cdot \mathbf{y}$ v $\mathbb{R}^n$.

2. Matice $\mathbf{G} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ je positivně definitní podle předchozí věty: $\mathbf{G}$ je symetrická a platí nerovnosti $\det(\mathbf{G}_1) = \det(1) > 0$ a $\det(\mathbf{G}_2) = \det(\mathbf{G}) > 0$.

   Připomeňme: $\mathbf{G}$ zadává „nezvyklý“ skalární součin $\langle \mathbf{x} \mid \mathbf{y} \rangle = \mathbf{x}^T \cdot \mathbf{G} \cdot \mathbf{y}$ v $\mathbb{R}^2$.

Příklady (pokrač.)

3. Matice
$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
není positivně definitní podle předchozí věty: platí
$\det(\mathbf{G}_1) > 0$, $\det(\mathbf{G}_2) < 0$, $\det(\mathbf{G}_3) > 0$, $\det(\mathbf{G}_4) < 0$.

Připomeňme (minulá přednáška): $\mathbf{G}$ zadává „skalární součin“ $\langle \mathbf{x} \mid \mathbf{y} \rangle = \mathbf{x}^T \cdot \mathbf{G} \cdot \mathbf{y}$ v Minkowského časoprostoru $\mathbb{R}^4$.

Věta (obecný tvar skalárního součinu v $\mathbb{R}^n$)

1. Ať $\mathbf{G}$ je positivně definitní matice typu $n \times n$ nad $\mathbb{R}$. Potom maticový součin

$$\mathbf{x}^T \cdot \mathbf{G} \cdot \mathbf{y}$$

definuje skalární součin v $\mathbb{R}^n$.

2. Každý skalární součin $\langle - \mid - \rangle$ v $\mathbb{R}^n$ definuje positivně definitní[a] matici $\mathbf{G} = (g_{ij})_{i=1,\dots,n, j=1,\dots,n}$, kde $g_{ij} = \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle$. Potom platí rovnost $\langle \mathbf{x} \mid \mathbf{y} \rangle = \mathbf{x}^T \cdot \mathbf{G} \cdot \mathbf{y}$.

[a] Matici $\mathbf{G}$ říkáme metrický tensor (také: Gramova matice) skalárního součinu $\langle - \mid - \rangle$.

Důkaz.

Přednáška. ■

Příklad (popis všech skalárních součinů v $\mathbb{R}^2$)

Matice $\mathbf{G}$ rozměrů 2×2 nad $\mathbb{R}$ je positivně definitní právě tehdy, když platí:

1. $\mathbf{G}$ je symetrická matice.
2. $\det(\mathbf{G}_1) > 0$ a $\det(\mathbf{G}_2) = \det(\mathbf{G}) > 0$.

To znamená: $\mathbf{G}$ je positivně definitní právě tehdy, když

$$\mathbf{G} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{kde } a > 0 \text{ a } ac - b^2 > 0$$

Neboli: vzorec

$$ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2$$

zadává skalární součin $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle$ v $\mathbb{R}^2$ právě tehdy, když platí nerovnosti $a > 0$ a $ac - b^2 > 0$.

Příklad (jednotková kružnice pro skalární součin v $\mathbb{R}^2$)

Pro positivně definitní[a] matici $\mathbf{G} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ a příslušný skalární součin $\langle - \mid - \rangle$ je množina[b]

$$S_1 = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid d(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}) = 1 \}$$

jednotková kružnice se středem v počátku.

Protože[c] $d(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}) = \| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \| = \| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \|$ a protože $\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \| = 1$ právě tehdy, když $\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \|^2 = 1$, platí

$$S_1 = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid a x_1^2 + 2b x_1 x_2 + c x_2^2 = 1 \}$$

a my ukážeme, že S_1 je elipsa.

---

[a] Připomenutí: platí $a > 0$ a $ac - b^2 > 0$.

[b] Připomenutí: d je metrika vytvořená skalárním součinem $\langle - \mid - \rangle$.

[c] Připomenutí: $\| - \|$ je norma vytvořená skalárním součinem $\langle - \mid - \rangle$.

Příklad (jednotková kružnice, pokrač.)

Například úpravou na čtverec je

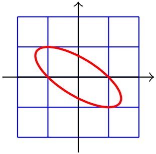
$$S_1 = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid a(x_1 + \frac{b}{a}x_2)^2 + \frac{ac - b^2}{a}x_2^2 = 1\}$$

Protože platí $a > 0$ a $ac - b^2 > 0$, je S_1 opravdu elipsa.

Například pro $\mathbf{G} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ je

$$S_1 = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid (x_1 + x_2)^2 + x_2^2 = 1\}$$

s (očekávaným) grafem



Připomenutí

Minulá přednáška: každý skalární součin vytváří normu.
Je-li $\langle - \mid - \rangle$ skalární součin na $\mathbb{R}^n$, pak

1. vektory $\mathbf{x}$ a $\mathbf{y}$ jsou ortogonální (také: navzájem kolmé), pokud $\langle \mathbf{x} \mid \mathbf{y} \rangle = 0$,
2. norma (také: velikost) vektoru $\mathbf{x}$ je $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x} \rangle}$,
3. vektor $\mathbf{x}$ je normovaný, pokud $\|\mathbf{x}\| = 1$.

Tvrzení (kanonická báze $\mathbb{R}^n$ a standardní skalární součin v $\mathbb{R}^n$)

Pro standardní skalární součin $\langle \mathbf{x} \mid \mathbf{y} \rangle = \mathbf{x}^T \cdot \mathbf{y}$ a kanonickou bázi $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ v $\mathbb{R}^n$ platí:[a] $\langle \mathbf{e}_i \mid \mathbf{e}_j \rangle = \mathbf{e}_i^T \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 0, & \text{pro } i \neq j \\ 1, & \text{pro } i = j \end{cases}$.

---

[a] Takovým bázím budeme říkat ortonormální a obecně je budeme studovat příště. To jest: vektory takové báze jsou na sebe navzájem kolmé a každý vektor takové báze má normu 1.

Věta (každou bázi $\mathbb{R}^n$ lze považovat za ortonormální)

Ať $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ je jakákoli uspořádaná báze $\mathbb{R}^n$. Potom existuje jediný skalární součin $\langle - \mid - \rangle$ takový, že $\langle \mathbf{b}_i \mid \mathbf{b}_j \rangle = \delta_{ij}$.

Důkaz.

Označme $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$, připomenutí (téma 05B):
$\mathbf{T}_{B \mapsto K_n} \cdot \mathbf{e}_i = \mathbf{b}_i$, čili $\mathbf{T}_{K_n \mapsto B} \cdot \mathbf{b}_i = \mathbf{e}_i$, pro vš $i = 1, \ldots, n$.

1. Existence hledaného skalárního součinu.
   Definujte $\mathbf{G} = (\mathbf{T}_{K_n \mapsto B})^T \cdot \mathbf{T}_{K_n \mapsto B}$. Potom matice $\mathbf{G}$ je positivně definitní a platí

$$\begin{aligned} \langle \mathbf{b}_i \mid \mathbf{b}_j \rangle &= \mathbf{b}_i^T \cdot \mathbf{G} \cdot \mathbf{b}_j \\ &= \underbrace{\mathbf{b}_i^T \cdot (\mathbf{T}_{K_n \mapsto B})^T}_{=(\mathbf{T}_{K_n \mapsto B} \cdot \mathbf{b}_i)^T = \mathbf{e}_i^T} \cdot \underbrace{\mathbf{T}_{K_n \mapsto B} \cdot \mathbf{b}_j}_{=\mathbf{e}_j} \\ &= \mathbf{e}_i^T \cdot \mathbf{e}_j = \delta_{ij} \end{aligned}$$

Důkaz (pokrač.)

2. Jednoznačnost hledaného skalárního součinu.
   Ať $\mathbf{b}_i^T \cdot \mathbf{G}_1 \cdot \mathbf{b}_j = \mathbf{b}_i^T \cdot \mathbf{G}_2 \cdot \mathbf{b}_j = \delta_{ij}$. Ukážeme $\mathbf{G}_1 = \mathbf{G}_2$.

   Opravdu: platí $\mathbf{b}_i^T \cdot (\mathbf{G}_1 - \mathbf{G}_2) \cdot \mathbf{b}_j = 0$ pro vš. i, j.

   To znamená $(\mathbf{T}_{B \mapsto K_n} \cdot \mathbf{e}_i)^T \cdot (\mathbf{G}_1 - \mathbf{G}_2) \cdot (\mathbf{T}_{B \mapsto K_n} \cdot \mathbf{e}_j) = 0$ pro vš. i, j, neboli $\mathbf{e}_i^T \cdot (\mathbf{T}_{B \mapsto K_n})^T \cdot (\mathbf{G}_1 - \mathbf{G}_2) \cdot \mathbf{T}_{B \mapsto K_n} \cdot \mathbf{e}_j = 0$ pro vš. i, j.

   Ukázali jsme rovnost $(\mathbf{T}_{B \mapsto K_n})^T \cdot (\mathbf{G}_1 - \mathbf{G}_2) \cdot \mathbf{T}_{B \mapsto K_n} = \mathbf{O}_{n,n}$.

   Protože $\mathbf{T}_{B \mapsto K_n}$ i $(\mathbf{T}_{B \mapsto K_n})^T$ jsou regulární, platí $\mathbf{G}_1 - \mathbf{G}_2 = \mathbf{O}_{n,n}$.

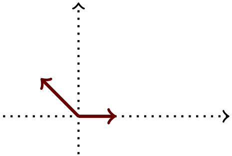
   Tudíž $\mathbf{G}_1 = \mathbf{G}_2$.

■

Příklad

Najděte skalární součin v $\mathbb{R}^2$ takový, aby vektory $\begin{pmatrix}1\\0\end{pmatrix}$ a $\begin{pmatrix}-1\\1\end{pmatrix}$ byly navzájem kolmé a každý měl normu 1.

Obrázek:



Skalární součin nalezneme podle předchozího tvrzení:

1. Pro $\mathbf{T}_{B\mapsto K_n} = \begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$ je[a] $(\mathbf{T}_{B\mapsto K_n})^{-1} = \begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$. Tudíž

   $$\mathbf{G} = ((\mathbf{T}_{B\mapsto K_n})^{-1})^T \cdot (\mathbf{T}_{B\mapsto K_n})^{-1} = \begin{pmatrix}1 & 1\\1 & 2\end{pmatrix}.$$

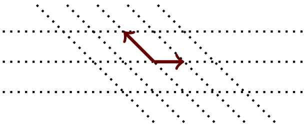
[a] Inversi matice $\mathbf{T}_{B\mapsto K_n}$ nalezneme nejrychleji pomocí adjungované matice.

Příklad (pokrač.)

2. Hledaný skalární součin je
$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 + x_1 y_2 + x_2 y_1 + 2 \cdot x_2 y_2.$$

3. Lze snadno spočítat: $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle = 0$, $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = 1$, $\left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \mid \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle = 1$.

Nalezený skalární součin „vidí“



jako jednotkovou pravoúhlou síť.

Co zatím v $\mathbb{R}^n$ umíme

Pro zadanou bázi $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ prostoru $\mathbb{R}^n$ umíme sestrojit skalární součin $\langle - \mid - \rangle$ v $\mathbb{R}^n$ tak, že $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ je ortonormální báze vzhledem k $\langle - \mid - \rangle$.

Příště se v $\mathbb{R}^n$ naučíme

Pro zadanou bázi $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ prostoru $\mathbb{R}^n$ a zadaný skalární součin $\langle - \mid - \rangle$ v $\mathbb{R}^n$ nalezneme novou bázi $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$, která je ortonormální vzhledem k $\langle - \mid - \rangle$.

Hledaná báze bude navíc splňovat rovnost $\mathrm{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k) = \mathrm{span}(\mathbf{c}_1, \ldots, \mathbf{c}_k)$ pro všechna $k = 1, \ldots, n$.

K tomu bude zapotřebí zavedení nových pojmů: ortogonální projekce a ortogonální rejekce.