# Cvičení 14

Nejprve trocha opakování z minula: Máme-li lineární prostor se skalárním součinem $L$ a uvnitř podprostor $W \subset L$ s ortogonální bází $(\vec{y}_1, \ldots, \vec{y}_k)$, pak lze snadno spočítat ortogonální průmět libovolného vektoru $\vec{v} \in L$ do $W$ jako

$$\operatorname{proj}_W \vec{v} = \frac{\langle \vec{y}_1 \,|\, \vec{v} \rangle}{\|\vec{y}_1\|^2} \vec{y}_1 + \frac{\langle \vec{y}_2 \,|\, \vec{v} \rangle}{\|\vec{y}_2\|^2} \vec{y}_2 + \cdots + \frac{\langle \vec{y}_k \,|\, \vec{y} \rangle}{\|\vec{y}_k\|^2} \vec{y}_k.$$

Nyní si zopakujeme z přednášky proces zvaný **Gramova–Schmidtova ortogonalizace**. Mějme nějaký lineární prostor $L$ se skalárním součinem (představujme si třeba nějaký podprostor $L \subset \mathbb{R}^m$) a chceme najít jeho ortogonální (či ortonormální) bázi. Začneme s nějakou bází $\mathscr{X} = (\vec{x}_1, \ldots, \vec{x}_n)$ prostoru $L$ a budeme sestavovat ortogonální bázi $\mathscr{Y} = (\vec{y}_1, \ldots, \vec{y}_n)$.

Začneme tím, že vezmeme $\vec{y}_1 := \vec{x}_1$. Dál bychom chtěli do báze přidat i vektor $\vec{x}_2$. To ale nemůžeme jen tak, neboť $\vec{x}_2$ nemusí být kolmý na $\vec{x}_1$. Uděláme to tedy tak, že od vektoru $\vec{x}_2$ odečteme jeho kolmý průmět do směru vektoru $\vec{y}_1 = \vec{x}_1$. Získám tak vektor

$$\vec{y}_2 := \vec{x}_2 - \frac{\langle \vec{y}_1 \,|\, \vec{x}_2 \rangle}{\|\vec{y}_1\|^2} \vec{y}_1,$$

jenž *je* kolmý na vektor $\vec{y}_1$.

Takto pokračujeme dále a dále. V $i$-tém kroku definujeme vektor

$$\vec{y}_i := \vec{x}_i - \frac{\langle \vec{y}_1 \,|\, \vec{x}_i \rangle}{\|\vec{y}_1\|^2} \vec{y}_1 - \cdots - \frac{\langle \vec{y}_{i-1} \,|\, \vec{x}_i \rangle}{\|\vec{y}_{i-1}\|^2} \vec{y}_{i-1}.$$

**Cvičení 14.1.** Uvažujme v prostoru $\mathbb{R}^3$ se standardním skalárním součinem rovinu $V = \operatorname{span}\{\vec{x}_1, \vec{x}_2\}$, kde

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \vec{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Nalezněte ortogonální bázi prostoru $V$ a doplňte ji na ortogonální bázi $\mathbb{R}^3$. S využitím tohoto výpočty najděte ortogonální průmět vektoru $\vec{v} = \left(\begin{smallmatrix} 1 \\ -1 \\ -3 \end{smallmatrix}\right)$ do roviny $V$.

Dále si uděláme krátký výlet do světa nestandardních skalárních součinů.

**Definice/Věta.** Matice $\mathbf{G}\colon \mathbb{R}^n \to \mathbb{R}^n$ se nazývá **pozitivně definitní**, jestliže je splněna některá z následujících ekvivalentních podmínek

(i) $\mathbf{G}$ je symetrická a $\vec{x}^{\mathrm{T}} \mathbf{G} \vec{x} > 0$ pro každé $\vec{x} \neq \vec{o}$ (a tedy $\langle \vec{x} \,|\, \vec{y} \rangle := \vec{x}^{\mathrm{T}} \mathbf{G} \vec{y}$ definuje skalární součin v $\mathbb{R}^n$)
(ii) $\mathbf{G}$ je symetrická a má všechna vlastní čísla kladná,
(iii) existuje matice s lineárně nezávislými sloupci $\mathbf{R}\colon \mathbb{R}^n \to \mathbb{R}^m$ taková, že $\mathbf{G} = \mathbf{R}^{\mathrm{T}} \mathbf{R}$,
(iv) existuje regulární matice $\mathbf{R}\colon \mathbb{R}^n \to \mathbb{R}^n$ taková, že $\mathbf{G} = \mathbf{R}^{\mathrm{T}} \mathbf{R}$.

**Věta.** Každý skalární součin v $\mathbb{R}^n$ lze vyjádřit ve tvaru $\langle \vec{x} \,|\, \vec{y} \rangle := \vec{x}^{\mathrm{T}} \mathbf{G} \vec{y}$, kde $\mathbf{G}$ je nějaká pozitivně definitní matice. Matice $\mathbf{G}$ se pak nazývá **metrický tenzor** nebo **Gramova matice** tohoto skalárního součinu.

**Poznámka.** Do teď jsme používali standardní skalární součin $\langle \vec{x} \,|\, \vec{y} \rangle = \vec{x}^{\mathrm{T}} \vec{y}$, jenž odpovídal metrickému tenzoru $\mathbf{G} = \mathbf{E}$.

**Věta.** (Sylvestrovo kritérium) Matice $\mathbf{G}$ je pozitivně definitní právě tehdy, když je symetrická a má všechny hlavní subdeterminanty kladné (tj. determinanty všech podmatic, které tvoří libovolně velký levý horní roh matice $\mathbf{G}$).

**Cvičení 14.2.** Rozhodněte o pozitivní definitnosti následujících matic

a) $\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 2 & -1 & 3 \\ 1 & 5 & 4 \\ 3 & 4 & 3 \end{pmatrix}$

Nakonec se podíváme na **metodu nejmenších čtverců**.

**Problém.** Řešíme soustavu lineárních rovnic $\mathbf{A}\vec{x} = \vec{b}$, která je tzv. *přeurčená* (angl. *overdetermined*), tj. máme víc rovnic než neznámých. Představujme si, že koeficienty v soustavě jsou určené nějakým fyzikálním měřením a neznámé jsou nějaké fyzikální parametry. Taková soustava nebude mít skoro jistě žádné řešení – fyzikální měření děláme s chybou, a tak spolu rovnice nebudou kompatibilní. Důležité ale je, že z nějakého důvodu víme, že ta soustava není úplně nesmyslná. Řešení by z fyzikálních důvodů mělo existovat, jen jsme ty naše měření neudělali dost přesně.

Ten nejstupidnější způsob jak náš problém vyřešit by byl zahodit nějaké rovnice tak, abychom měli stejný počet rovnic jak neznámých. To by jistě k nějakému výsledku vedlo, ale zbytečně bychom se připravili o drahocenně naměřená data. Hledáme nějaký sofistikovaný způsob, jak najít to nejlepší možné řešení, které se co nejvíc blíží realitě. Zkusme najít nějakou matematicky přesnou charakterizaci: Budeme hledat takové řešení $\vec{x}$, že $\|\mathbf{A}\vec{x} - \vec{b}\|$ je minimální.

**Návod.** Řešme soustavu $\mathbf{A}^{\mathrm{T}}\mathbf{A}\vec{x} = \mathbf{A}^{\mathrm{T}}\vec{b}$. (Matice $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ je určitě čtvercová, takže teď už máme stejně rovnic jako neznámých,)

**Zdůvodnění.** Chceme minimalzovat $\|\mathbf{A}\vec{x} - \vec{b}\|$, tj. hledáme bod $\vec{y} = \mathbf{A}\vec{x} \in \operatorname{im}\mathbf{A}$ takový, který je „co nejblíže“ bodu $\vec{b}$. Ten se dá najít tak, že spočítáme kolmý průmět $\vec{b}$ do $\operatorname{im}\mathbf{A}$, tj.

$$\mathbf{A}\vec{x} = \vec{y} = \operatorname{proj}_{\operatorname{im} A}\vec{b} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\vec{b},$$

vynásobením $\mathbf{A}^{\mathrm{T}}$ zleva dostaneme

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\vec{x} = \mathbf{A}^{\mathrm{T}}\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\vec{b} = \mathbf{A}^{\mathrm{T}}\vec{b}.$$

**Cvičení 14.3.** Metodou nejmenších čtverců vyřešte následující soustavy. Pro každé řešení vypočítejte odchylku $\|\mathbf{A}\vec{x} - \vec{b}\|$.

a) $\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{array}\right)$ b) $\left(\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 3 \end{array}\right)$

**Cvičení 14.4.** Kvadratickou funkcí $f(x) = a + bx + cx^2$ proložte následující body

| $x$ | -1 | 0 | 1 | 2 |
|---|---|---|---|---|
| $y$ | 3 | 1 | -1 | 1 |

Lineární regresi jsem naprogramoval v SAGE zde[1]. Po kliknutí na odkaz budete nejspíš chvilku muset nějakou dobu čekat a pak se vám otevře interaktivní sešit, kde je vyřešené výše uvedené cvičení. Zadání si můžete upravit podle sebe a s výsledky si hrát.

# Řešení

1. OG báze třeba $\left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}\right)$, $\operatorname{proj}_V \vec{v} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$
2. ano, ne, ne
3. $\binom{2}{0}$, řešení přesné; $\binom{1/2}{5/3}$, odchylka $\|\mathbf{A}\vec{x} - \vec{b}\|^2 = 1/6$
4. $f(x) = x^2 - 9/5\,x + 2/5$

[1] https://mybinder.org/v2/gh/gromadan/LAG_sage/master?filepath=regrese.ipynb