Komentář k oddílu IX.6 - stopa matice

Definice stopy a Věta IX.21:

- Stopa se definuje pro čtvercové matice.

  Je definována jako součet prvků na hlavní diagonále.

  Stopa matice $\mathbb{A}$ se značí $\operatorname{tr}(\mathbb{A})$ (z anglického názvu *trace*).

  V teorii matic hraje stopa důležitou roli, ale my to zas tak moc neoceníme. Hlavní motivací pro nás je Věta IX.22, která se snad používá v ekonometrii.

- Věta IX.21 shrnuje základní vlastnosti stopy. Body (i) a (ii) jsou zřejmé:

  Matice $\mathbb{A} + \mathbb{B}$ má na diagonále prvky $a_{ii} + b_{ii}$, takže jejich součet je roven součtu stop matic $\mathbb{A}$ a $\mathbb{B}$.

  Matice $\alpha\mathbb{A}$ má na diagonále prvky αa_{ii}, takže jejich součet je roven α-násobku stopy matice $\mathbb{A}$.

- Bod (iii) je klíčovou vlastností stopy. Násobení matic není komutativní, ale součiny $\mathbb{A}\mathbb{B}$ a $\mathbb{B}\mathbb{A}$ mají stejnou stopu.

  To není vidět na první pohled, ale dokáže se to přímým výpočtem:

  Matice $\mathbb{A}\mathbb{B}$ má na místě ii číslo

  $$\sum_{j=1}^{n} a_{ij} b_{ji},$$

  tedy

  $$\operatorname{tr}(\mathbb{A}\mathbb{B}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji}.$$

  Matice $\mathbb{B}\mathbb{A}$ má na místě ii číslo

  $$\sum_{j=1}^{n} b_{ij} a_{ji},$$

  tedy

  $$\operatorname{tr}(\mathbb{B}\mathbb{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} a_{ji}.$$

Porovnáme-li výsledky, vidíme, že se shodují: Násobení čísel je komutativní, a tedy $b_{ij}a_{ji} = a_{ji}b_{ij}$. Zároveň je sčítání čísel komutativní a asociativní, takže nezáleží na pořadí, v jakém sčítáme.

Oba výsledky se rovnají součtu všech součinů tvaru $a_{ij}b_{ji}$, kde $i, j \in \{1, \ldots, n\}$. (Pouze se sčítají v jiném pořadí.)

- Bod (iv) je důsledkem bodu (iii):

$$\operatorname{tr}(\mathbb{ABC}) = \operatorname{tr}((\mathbb{AB})\mathbb{C}) \overset{(iii)}{=} \operatorname{tr}(\mathbb{C}(\mathbb{AB})) = \operatorname{tr}((\mathbb{CA})\mathbb{B}) \overset{(iii)}{=} \operatorname{tr}(\mathbb{B}(\mathbb{CA})) .$$

- Není pravda, že by stopa součinu matic nezávisela na jejich pořadí. To platí jen pro dvojici matic. Bod (iv) sice mluví o trojicích, ale prohazují se vždy jen dvojice matic, jak je zřejmé z důkazu.

  Například se tedy může stát, že

$$\operatorname{tr}(\mathbb{ABC}) \neq \operatorname{tr}(\mathbb{CBA}) .$$

  Svědčí o tom například volba

$$\mathbb{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \mathbb{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \ \mathbb{C} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

  Pak totiž

$$\begin{aligned} \operatorname{tr}(\mathbb{ABC}) &= \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) = \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ -4 & 3 \end{pmatrix}\right) \\ &= \operatorname{tr}\left(\begin{pmatrix} -4 & 3 \\ -2 & 1 \end{pmatrix}\right) = -4 + 1 = -3; \\ \operatorname{tr}(\mathbb{CBA}) &= \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \operatorname{tr}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}\right) \\ &= \operatorname{tr}\left(\begin{pmatrix} 4 & 3 \\ -2 & -1 \end{pmatrix}\right) = 4 - 1 = 3. \end{aligned}$$

- Stopa matice se rovná součtu vlastních čísel, pokud každé počítáme tolikrát, kolik je jeho násobnost. To už jsme poznamenali v poznámce za Větou IX.18, ale podrobný důkaz jsme nedělali.

  Pro symetrické matice toto tvrzení plyne z Vět IX.20 a IX.21.

Je-li $\mathbb{A}$ symetrická matice, pak podle Věty IX.20 existuje diagonální matice $\mathbb{D}$ (na diagonále má vlastní čísla matice $\mathbb{A}$) a ortogonální matice $\mathbb{Q}$, že $\mathbb{A} = \mathbb{Q}\mathbb{D}\mathbb{Q}^T$.

Pak

$$\mathrm{tr}\,(\mathbb{A}) = \mathrm{tr}\,\left(\mathbb{Q}\mathbb{D}\mathbb{Q}^T\right) \overset{IX.21(iv)}{=} \mathrm{tr}\,\left(\underbrace{\mathbb{Q}^T\mathbb{Q}}_{=\mathbb{I}}\,\mathbb{D}\right) = \mathrm{tr}\,(\mathbb{D})\,.$$

Idempotetní matice a Věta IX.22:

- Čtvercová matice $\mathbb{A}$ řádu n se nazývá idempotentní, pokud $\mathbb{A}^2 = \mathbb{A}$.

  To znamená, že lineární zobrazení reprezentované maticí $\mathbb{A}$ je *projekce*. Vysvětleme to:

  Nechť $L : \mathbf{R}^n \to \mathbf{R}^n$ je zobrazení reprezentované idempotentní maticí $\mathbb{A}$. Toto zobrazení je definováno vzorcem

  $$L(\boldsymbol{x}) = \mathbb{A}\boldsymbol{x}, \qquad \boldsymbol{x} \in \mathbf{R}^n.$$

  Připomeňme, že $\mathrm{Im}\, L$ je obor hodnot zobrazení L a že to podprostor prostoru $\mathbf{R}^n$ (viz Věta IX.5).

  Nechť $\boldsymbol{x} \in \mathrm{Im}\, L$. To znamená, že existuje $\boldsymbol{y} \in \mathbf{R}^n$ splňující $L(\boldsymbol{y}) = \boldsymbol{x}$. Pak

  $$L(\boldsymbol{x}) = \mathbb{A}\boldsymbol{x} = \mathbb{A} \cdot L(\boldsymbol{y}) = \mathbb{A} \cdot \mathbb{A}\boldsymbol{y} = \mathbb{A}^2\boldsymbol{y} = \mathbb{A}\boldsymbol{y} = L(\boldsymbol{y}) = \boldsymbol{x}.$$

  Použili jsme, že L je reprezentováno maticí $\mathbb{A}$; skutečnost, že $L(\boldsymbol{y}) = \boldsymbol{x}$, a navíc předpoklad, že $\mathbb{A}$ je idempotentní (v páté rovnosti).

  Tedy,

  $$\text{pro každé } \boldsymbol{x} \in \mathrm{Im}\, L \text{ platí } L(\boldsymbol{x}) = \boldsymbol{x}. \tag{$*$}$$

  Dále, pro každé $\boldsymbol{x} \in \mathbf{R}^n$ platí

  $$\boldsymbol{x} = \underbrace{L(\boldsymbol{x})}_{\in \mathrm{Im}\, L} + (\underbrace{\boldsymbol{x} - L(\boldsymbol{x})}_{\in \ker L})$$

  Vektor $\boldsymbol{x} - L(\boldsymbol{x})$ patří do jádra L, protože

  $$L(\boldsymbol{x} - L(\boldsymbol{x})) = L(\boldsymbol{x}) - L(\underbrace{L(\boldsymbol{x})}_{\in \mathrm{Im}\, L}) \overset{(*)}{=} L(\boldsymbol{x}) - L(\boldsymbol{x}) = \boldsymbol{o}$$

Proto každé $\boldsymbol{x} \in \mathbf{R}^n$ lze právě jedním způsobem vyjádřit ve tvaru

$$\boldsymbol{x} = \underbrace{\boldsymbol{x}_1}_{\in \operatorname{Im} L} + \underbrace{\boldsymbol{x}_2}_{\in \ker L} \tag{$**$}$$

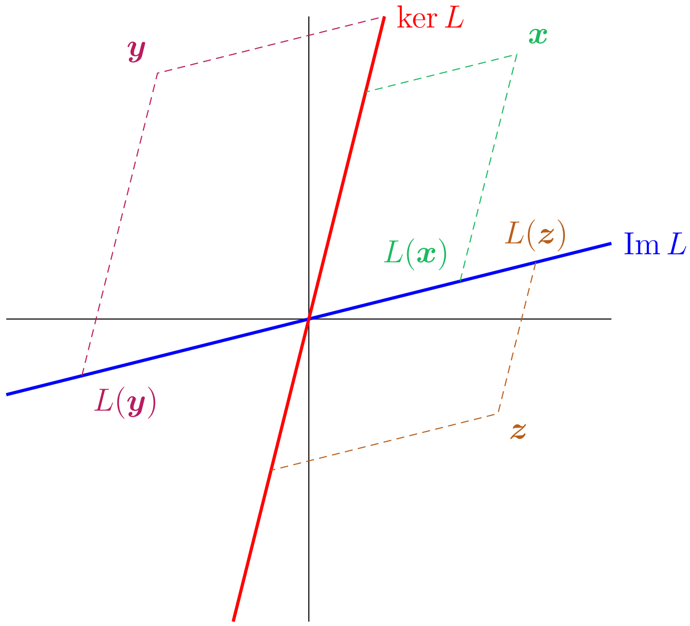
Pode výpočtu výše lze zvolit $\boldsymbol{x}_1 = L(\boldsymbol{x})$ a $\boldsymbol{x}_2 = \boldsymbol{x} - L(\boldsymbol{x})$. Navíc je to jediná možná volba. Pokud totiž $\boldsymbol{x}$ je vyjádřeno jako v $(**)$, pak

$$L(\boldsymbol{x}) = L(\boldsymbol{x}_1 + \boldsymbol{x}_2) = L(\underbrace{\boldsymbol{x}_1}_{\in \operatorname{Im} L}) + \underbrace{L(\boldsymbol{x}_2)}_{=\boldsymbol{o}} \overset{(*)}{=} \boldsymbol{x}_1 + \boldsymbol{o} = \boldsymbol{x}_1.$$

Pak ovšem nutně $\boldsymbol{x}_2 = \boldsymbol{x} - \boldsymbol{x}_1 = \boldsymbol{x} - L(\boldsymbol{x})$.

Tedy, každý vektor $\boldsymbol{x}$ se dá právě jedním způsobem vyjádřit ve tvaru $(**)$ a zobrazení L je definováno tak, že vektoru $\boldsymbol{x}$ se přiřadí první sčítanec $\boldsymbol{x}_1$.

Takovým zobrazením se říká projekce. Na následujícím obrázku je znázorněna jedna taková projekce v $\mathbf{R}^2$.



- Důkaz bodu (i): Nechť λ je vlastní číslo a $\boldsymbol{x}$ je k němu příslušný vlastní vektor. Pak platí:

$$\lambda \boldsymbol{x} = \mathbb{A}\boldsymbol{x} = \mathbb{A}^2\boldsymbol{x} = \mathbb{A}(\mathbb{A}\boldsymbol{x}) = \mathbb{A} \cdot \lambda \boldsymbol{x} = \lambda \cdot \mathbb{A}\boldsymbol{x} = \lambda \cdot \lambda \boldsymbol{x} = \lambda^2 \boldsymbol{x}.$$

Ve druhé rovnosti jsme použili, že $\mathbb{A}$ je idempotentní. Platí tedy $\lambda \boldsymbol{x} = \lambda^2 \boldsymbol{x}$, neboli $(\lambda - \lambda^2)\boldsymbol{x} = \boldsymbol{o}$. Protože $\boldsymbol{x} \neq \boldsymbol{o}$, je $\lambda - \lambda^2 = 0$, tedy $\lambda = 0$ nebo $\lambda = 1$.

- Bod (ii) plyne z bodu (i) a z Důsledku Věty IX.20.
- Důkaz bodu (iii):

  Popíšeme jak najít matici $\mathbb{Q}$.

  Nechť L je lineární zobrazení reprezentované maticí $\mathbb{A}$. Podle Věty IX.5 víme, že $\operatorname{Im} L$ je podprostor prostoru $\mathbf{R}^n$.

  Nechť $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ je nějaká báze $\operatorname{Im} L$ (existuje díky Větě IX.3, navíc $k \leq n$).

  Podle Věty IX.5 dále víme, že $\ker L$ je také podprostor prostoru $\mathbf{R}^n$. Nechť $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ je nějaká báze $\ker L$ (existuje díky Větě IX.3, navíc $l \leq n$).

  Věta IX.5 nakonec ještě dává rovnost

  $$\dim \mathbf{R}^n = \dim \operatorname{Im} L + \dim \ker L,$$

  tj.

  $$n = k + l.$$

  Výše jsme ukázali, že každé $\boldsymbol{x} \in \mathbf{R}^n$ lze zapsat ve tvaru $(**)$, tedy

  $$\operatorname{lin}_{\mathbf{R}}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_l\} = \mathbf{R}^n.$$

  (Každý prvek $\operatorname{Im} L$ lze vyjádřit jako lineární kombinaci vektorů $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$. Každý prvek $\ker L$ lze vyjádřit jako lineární kombinaci vektorů $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$. Každý prvek $\mathbf{R}^n$ lze vyjádřit jako součet prvku $\operatorname{Im} L$ a prvku $\ker L$. Tedy, každý prvek $\mathbf{R}^n$ lze vyjádřit jako lineární kombinaci vektorů $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$.)

  Protože $k+l = n = \dim \mathbf{R}^n$, Větička IX.4(ii) říká, že vektory $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ tvoří bázi $\mathbf{R}^n$. Speciálně jsou lineárně nezávislé.

  Nechť $\mathbb{Q}$ je matice typu $n \times n$, jejíž sloupce jsou $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$. Protože jsou tyto vektory lineárně nezávislé, je matice $\mathbb{Q}$ regulární.

  Protože vektory $\boldsymbol{v}_j$ patří do $\Im L$, platí

  $$\mathbb{A}\boldsymbol{v}_j = L(\boldsymbol{v}_j) = \boldsymbol{v}_j, \quad j = 1, \ldots, k.$$

Protože vektory $\boldsymbol{u}_j$ patří do $\ker L$, platí

$$\mathbb{A}\boldsymbol{u}_j = L(\boldsymbol{u}_j) = \boldsymbol{o}, \quad j = 1, \ldots, l.$$

Tedy

$$\begin{aligned}
\mathbb{A}\mathbb{Q} &= \mathbb{A}\begin{pmatrix}\boldsymbol{v}_1 & \ldots & \boldsymbol{v}_k & \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_l\end{pmatrix} = \begin{pmatrix}\mathbb{A}\boldsymbol{v}_1 & \ldots & \mathbb{A}\boldsymbol{v}_k & \mathbb{A}\boldsymbol{u}_1 & \ldots & \mathbb{A}\boldsymbol{u}_l\end{pmatrix} \\
&= \begin{pmatrix}\boldsymbol{v}_1 & \ldots & \boldsymbol{v}_k & \boldsymbol{o} & \ldots & \boldsymbol{o}\end{pmatrix} \\
&= \left(\begin{array}{cccc|ccc}
1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\
\hline
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{array}\right)\begin{pmatrix}\boldsymbol{v}_1 & \ldots & \boldsymbol{v}_k & \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_l\end{pmatrix} \\
&= \left(\begin{array}{cccc|ccc}
1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\
\hline
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{array}\right)\mathbb{Q},
\end{aligned}$$

kde uvedená diagonální matice má na diagonále k jedniček $l = n - k$ nul. Protože $\mathbb{Q}$ je regulární, dostaneme

$$\mathbb{Q}\mathbb{A}\mathbb{Q}^{-1} = \left(\begin{array}{cccc|ccc}
1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\
\hline
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{array}\right),$$

což dokončuje důkaz bodu (iii).

- Důkaz bodu (iv):

  Platí
  $$\operatorname{tr}\left(\mathbb{Q}\mathbb{A}\mathbb{Q}^{-1}\right) \overset{IX.21(iv)}{=} \operatorname{tr}\left(\underbrace{\mathbb{Q}^{-1}\mathbb{Q}}_{=\mathbb{I}}\mathbb{A}\right) = \operatorname{tr}(\mathbb{A}).$$
  Navíc zřejmě
  $$\operatorname{tr}\left(\mathbb{Q}\mathbb{A}\mathbb{Q}^{-1}\right) = k = h(\mathbb{Q}\mathbb{A}\mathbb{Q}^{-1}),$$
  protože $\mathbb{Q}\mathbb{A}\mathbb{Q}^{-1}$ je schodovitá matice.

  Nakonec,
  $$k = \dim \operatorname{Im} L = h(\mathbb{A}).$$
  První rovnost plyne z konstrukce v důkaze bodu (iii), druhá byla vysvětlena v závěru komentáře k oddílu IX.2.

- Důkaz bodu (v):

  Označme $H : \mathbf{R}^n \to \mathbf{R}^n$ lineární zobrazení reprezentované maticí $\mathbb{I} - \mathbb{A}$. Tj.,
  $$H(\boldsymbol{x}) = \boldsymbol{x} - \mathbb{A}\boldsymbol{x} = \boldsymbol{x} - L(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbf{R}^n.$$
  Připomeňme, že každé $\boldsymbol{x} \in \mathbf{R}^n$ lze jednoznačně vyjádřit ve tvaru $(**)$. Pak
  $$H(\boldsymbol{x}) = H(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{x}_1 + \boldsymbol{x}_2 - \underbrace{L(\boldsymbol{x}_1 + \boldsymbol{x}_2)}_{=\boldsymbol{x}_1} = \boldsymbol{x}_2$$
  Tedy
  $$\ker H = \operatorname{Im} L \text{ a } \operatorname{Im} H = \ker L.$$
  Z toho vidíme, že
  $$h(\mathbb{I} - \mathbb{A}) = \dim \operatorname{Im} H = \dim \ker L = l = n - k = n - h(\mathbb{A}).$$

- Jiný důkaz bodu (v):

  Zřejmě
  $$\operatorname{tr}(\mathbb{I} - \mathbb{A}) = \operatorname{tr}(\mathbb{I}) - \operatorname{tr}(\mathbb{A}) = n - k = n - h(\mathbb{A}).$$
  Dále, matice $\mathbb{I} - \mathbb{A}$ je také idempotentní, protože
  $$(\mathbb{I} - \mathbb{A})^2 = (\mathbb{I} - \mathbb{A})(\mathbb{I} - \mathbb{A}) = \underbrace{\mathbb{I}^2}_{=\mathbb{I}} - \underbrace{\mathbb{I}\mathbb{A}}_{=\mathbb{A}} - \underbrace{\mathbb{A}\mathbb{I}}_{=\mathbb{A}} + \underbrace{\mathbb{A}^2}_{=\mathbb{A}} = \mathbb{I} - \mathbb{A}.$$
  Tedy z bodu (iv) aplikovaného na $\mathbb{I} - \mathbb{A}$ plyne
  $$h(\mathbb{I} - \mathbb{A}) = \operatorname{tr}(\mathbb{I} - \mathbb{A}) = n - h(\mathbb{A}).$$