TODO!!!

Definice 0.1 (Dot product on the space of matrices)

$$\mathbb{A} : \mathbb{B} = \mathrm{tr}(\mathbb{A}\mathbb{B}^T).$$

Definice 0.2 (Norm of matrix)

$$|\mathbb{A}| = (\mathbb{A} : \mathbb{A})^{\frac{1}{2}}.$$

Příklad

$$(\mathbf{a} \otimes \mathbf{b})^T = \mathbf{b} \otimes \mathbf{a}.$$

Důkaz

$$\mathbf{u} \cdot (\mathbf{a} \otimes \mathbf{b})^T \mathbf{v} = (\mathbf{a} \otimes \mathbf{b})\mathbf{u} \cdot \mathbf{v} = (\mathbf{a}(\mathbf{b} \cdot \mathbf{u}))\mathbf{v} = (\mathbf{b} \cdot \mathbf{u})(\mathbf{a} \cdot \mathbf{v}) = \mathbf{u} \cdot (\mathbf{b}(\mathbf{a} \cdot \mathbf{v})) = \mathbf{u} \cdot (\mathbf{b} \otimes \mathbf{a})\mathbf{v}.$$

□

Příklad

$$\det(e^{\mathbb{A}}) = e^{\mathrm{tr}\,\mathbb{A}}.$$

Důkaz

$$e^{\mathbb{A}} = \lim \left(\mathbb{I} + \frac{\mathbb{A}}{n}\right)^n.$$

$$\det e^{\mathbb{A}} = \lim_{n\to\infty} \left(\det\left(\mathbb{I} + \frac{\mathbb{A}}{n}\right)^n\right) = \lim_{n\to\infty} \left(\det\left(\mathbb{I} + \frac{\mathbb{A}}{n}\right)\right)^n =?$$

Subtask: Is there an approximation for $\det(\mathbb{I} + \mathbb{S})$, where $\mathbb{S}$ is a „small“ matrix. Yes, we did it (KontinuumDU1.pdf) for $\mathbb{S} \in \mathbb{R}^{3\times 3}$:

$$\det(\mathbb{I}+\mathbb{S}) = \det\mathbb{I}+\operatorname{tr}(\mathbb{I}\operatorname{cof}\mathbb{S})+\operatorname{tr}(\mathbb{S}^T\operatorname{cof}\mathbb{I})+\det\mathbb{S} \approx 1+\operatorname{tr}(\mathbb{S}^T\operatorname{cof}\mathbb{I})+o(\mathbb{S}^2) = 1+\operatorname{tr}(S)+o(\mathbb{S}^2).$$

And for $\mathbb{S} \in \mathbb{R}^{n\times n}$, one can see that:

$$\det(\mathbb{I}+\mathbb{S}) = \begin{pmatrix} 1+s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & 1+s_{22} & \dots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \dots & 1+s_{nn} \end{pmatrix} = (1+s_{11})(1+s_{22})\cdot\ldots\cdot(1+s_{nn})+o(\mathbb{S}^2) =$$

$$= 1 + s_{11} + s_{22} + \ldots + s_{nn} + o(\mathbb{S}^2) = 1 + \operatorname{tr}\mathbb{S} + o(\mathbb{S}^2).$$

$$? = \lim_{n\to\infty}\left(1 + \frac{\operatorname{tr}\mathbb{A}}{n} + \ldots\right)^n = e^{\operatorname{tr}\mathbb{A}}.$$

□

Tvrzení 0.1

$$\det(\mathbb{I} + \mathbb{S}) = 1 + \operatorname{tr}\mathbb{S} + \ldots$$

Definice 0.3 (Gateaux derivative)

$$D\mathbf{f}(\mathbf{x})[\mathbf{y}] := \frac{d}{d\tau}\mathbf{f}(\mathbf{x} + \tau\mathbf{y})|_{\tau=0}.$$

Definice 0.4 (Fréchet derivative)

$\mathbf{f} : U \to V$:

$$\lim_{\|\mathbf{y}\|_U\to 0} \frac{\|\mathbf{f}(\mathbf{x}+\mathbf{y}) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x})[\mathbf{y}]\|_V}{\|\mathbf{y}\|_V} = 0.$$

Poznámka

Sometimes we write $\nabla f(\mathbf{x}) \cdot \mathbf{y}$ instead of $Df(\mathbf{x})[\mathbf{y}]$ (from Riesz representation theorem).

For matrices ($\varphi : \mathbb{A} \in \mathbb{R}^{3\times 3} \to \mathbb{R}$):

$$\frac{\|\varphi(\mathbb{A} + \mathbb{B}) - \varphi(\mathbb{A}) - D\varphi(\mathbb{A})[\mathbb{B}]\|_{\mathbb{R}}}{\|\mathbb{B}\|_{\mathbb{R}^{3\times 3}}}.$$

Poznámka
We write $\frac{\partial\varphi}{\partial\mathbb{A}}(\mathbb{A}) : \mathbb{B}$ instead of $D\varphi(\mathbb{A})[\mathbb{B}]$, where $\frac{\partial\varphi}{\partial\mathbb{A}}(\mathbb{A})$ is right matrix. Warning $\frac{\partial\varphi}{\partial\mathbb{A}}(\mathbb{A}) \neq D\varphi(\mathbb{A})$, because of transposition ($\mathbb{A} : \mathbb{B} = \operatorname{tr}(\mathbb{A}\mathbb{B}^T) = \operatorname{tr}(\mathbb{A}^T\mathbb{B})$).

Příklad

$$\frac{\partial \operatorname{tr}\mathbb{A}}{\partial\mathbb{A}}(\mathbb{A})[\mathbb{B}] = \frac{d}{d\tau}(\operatorname{tr}(\mathbb{A}+\tau\mathbb{B}))|_{\tau=0} = \frac{d}{d\tau}\left(\operatorname{tr}\mathbb{A} + \tau\operatorname{tr}\mathbb{B}\right)|_{\tau=0} = \operatorname{tr}\mathbb{B} = \mathbb{I} : \mathbb{B}.$$

So $\frac{\partial \operatorname{tr}\mathbb{A}}{\partial\mathbb{A}} = \mathbb{I}$.

Příklad

$$\frac{\partial \det\mathbb{A}}{\partial\mathbb{A}}(\mathbb{A})[\mathbb{B}] = \frac{d}{d\tau}(\det(\mathbb{A}+\tau\mathbb{B}))|_{\tau=0} = \frac{d}{d\tau}\left(\det(\mathbb{A})\cdot\det\left(\mathbb{I}+\tau\mathbb{A}^{-1}\mathbb{B}\right)\right)|_{\tau=0} =$$

$$= \frac{d}{d\tau}\left((\det\mathbb{A})\cdot\left(1+\tau\operatorname{tr}(\mathbb{A}^{-1}\mathbb{B}) + o(\tau^2)\right)\right)|_{\tau=0} = (\det\mathbb{A})\operatorname{tr}\left(\mathbb{A}^{-1}\mathbb{B}\right) =$$

$$= (\det\mathbb{A})\operatorname{tr}\left(\left(\mathbb{A}^{-T}\right)^T\mathbb{B}\right) = \left((\det\mathbb{A})\mathbb{A}^{-T}\right) : \mathbb{B}.$$

So $\frac{\partial \det\mathbb{A}}{\partial\mathbb{A}} = (\det\mathbb{A})\mathbb{A}^{-T} = \operatorname{cof}(\mathbb{A})$.

Příklad
$\mathbb{A} : \mathbb{R} \to \mathbb{R}^{3\times 3}$.

$$\frac{d}{dt}(\det\mathbb{A}(t)) = (\det\mathbb{A}(t))\operatorname{tr}\left(\mathbb{A}(t)^{-1}\frac{d\mathbb{A}(t)}{dt}\right).$$

Příklad
$\mathbb{F} : \mathbb{A} \in \mathbb{R}^{3\times3} \to \mathbb{F}(\mathbb{A}) \in \mathbb{R}^{3\times3}$. $\mathbb{F}(\mathbb{A}) = \mathbb{A}^{-1}$. (We know $\frac{1}{1+x} = 1 - x + \ldots$.)

$$\frac{\partial\mathbb{F}(\mathbb{A})}{\partial\mathbb{A}}(\mathbb{A})[\mathbb{B}] = \frac{d}{d\tau}\left((\mathbb{A}+\tau\mathbb{B})^{-1}\right)|_{\tau=0} = \frac{d}{d\tau}\left(\left(\mathbb{A}\left(\mathbb{I}+\tau\mathbb{A}^{-1}\mathbb{B}\right)\right)^{-1}\right)|_{\tau=0} =$$

$$= \frac{d}{d\tau}\left(\left(\mathbb{I}+\tau\mathbb{A}^{-1}\mathbb{B}\right)^{-1}\mathbb{A}^{-1}\right)|_{\tau=0} = \frac{d}{d\tau}\left(\left(\mathbb{I}-\tau\mathbb{A}^{-1}\mathbb{B}+\ldots\right)\mathbb{A}^{-1}\right)|_{\tau=0} = -\mathbb{A}^{-1}\mathbb{B}\mathbb{A}^{-1}.$$

So we have $\frac{\partial(\mathbb{A}^{-1})_{ij}}{\partial(\mathbb{A})_{kl}}(\mathbb{B})_{kl}$.

From chain rule (but this is easily solvable by differentiating $\mathbb{A}^{-1}(t)\mathbb{A}(t) = \mathbb{I}$):

$$\frac{d}{dt}\left(\mathbb{A}^{-1}\right) = -\mathbb{A}^{-1}\frac{d\mathbb{A}}{dt}\mathbb{A}^{-1}.$$

Příklad
$\mathbb{F}(\mathbb{A}) = e^{\mathbb{A}}$

$$\frac{\partial e^{\mathbb{A}}}{\partial \mathbb{A}}[\mathbb{B}] = \frac{d}{d\tau}(e^{\mathbb{A}+\tau\mathbb{B}})|_{\tau=0} = \frac{d}{d\tau}\left(\mathbb{I} + \frac{\mathbb{A}+\tau\mathbb{B}}{1!} + \frac{(\mathbb{A}+\tau\mathbb{B})^2}{2!}\right)|_{\tau=0}.$$

Věta 0.2 (Daleckii–Krein)

$\mathbb{A}$ real symmetric matrix, $\mathbb{A} \in \mathbb{R}^{k\times k}$, $\mathbb{A} = \sum_{i=1}^{k} \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i$, where λ_i are eigenvalues and $\mathbf{v}_i$ are normalised orthogonal ($\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}$) eigenvectors.

f continuously differentiable real function defined on open set containing the spectrum of $\mathbb{A}$

$$\mathbb{F}(\mathbb{A}) := \sum_{i=1}^{k} f(\lambda_i)\mathbf{v}_i \otimes \mathbf{v}_i =: \sum_{i=1}^{k} f(\lambda_i)\mathbb{P}_i.$$

Then the formula for the Gateaux derivative of f at point $\mathbb{A}$ in direction $\mathbb{X}$ reads

$$D\mathbb{F}(\mathbb{A})[\mathbb{X}] = \frac{\partial \mathbb{F}}{\partial \mathbb{A}}[\mathbb{X}] = \sum_{i=1}^{k} \frac{df}{d\lambda}|_{\lambda=\lambda_i}\mathbb{P}_i\mathbb{X}\mathbb{P}_i + \sum_{i=1}^{k}\sum_{j=1,j\neq i}^{k} \frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j}\mathbb{P}_i\mathbb{X}\mathbb{P}_j.$$

Sometimes we write $D\mathbb{F}(\mathbb{A})[\mathbb{X}] = f^{[1]}(\mathbb{A}) \bigcirc \mathbb{X}$ (Schur product of matrices, it is point-wise multiplication). Then

$$[f^{[1]}(\mathbb{A})]_{ij} = \begin{cases} \frac{df}{d\lambda}|_{\lambda=\lambda_i}, & i = j, \\ \frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j}, & i \neq j. \end{cases}$$

Důkaz
No summation conventions, all sums are stated explicitly!

$$\mathbb{F}(\mathbb{A}) = \sum_{i=1}^{k} f(\lambda_i)\mathbf{v}_i \otimes \mathbf{v}_i =$$

$$= \sum_{i=1}^{k} f(\lambda_i(a_{11}, a_{12}, \ldots, a_{21}, \ldots))\mathbf{v}_i(a_{11}, a_{12}, \ldots, a_{21}, \ldots) \otimes \mathbf{v}_i(a_{11}, a_{12}, \ldots, a_{21}, \ldots).$$

$$\frac{\partial \mathbb{F}(\mathbb{A})}{\partial \mathbb{A}} = \sum_{i=1}^{k} \left(\frac{\partial f}{\partial \lambda_i} \frac{\partial \lambda_i}{\partial \mathbb{A}} \mathbf{v}_i \otimes \mathbf{v}_i + f(\lambda_i) \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \otimes \mathbf{v}_i + f(\lambda_i)\mathbf{v}_i \otimes \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \right) =?.$$

We derivate $\mathbb{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$:

$$\frac{\partial \mathbb{A}}{\partial \mathbb{A}} \mathbf{v}_i + \mathbb{A} \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} = \frac{\partial \lambda_i}{\partial \mathbb{A}} \mathbf{v}_i + \lambda_i \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}}.$$

We multiply (with dot product) it by $\mathbf{v}_i$:

$$\mathbb{P}_i + \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \cdot \mathbb{A}^T \mathbf{v}_i = \frac{\partial \lambda_i}{\partial \mathbb{A}} \cdot 1 + \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \mathbb{A} \cdot \mathbf{v}_i.$$

$$\frac{\partial \lambda_i}{\partial \mathbb{A}} = \mathbb{P}_i = \mathbf{v}_i \otimes \mathbf{v}_i.$$

We again multiply derivative of $\mathbb{A}|\mathbf{v}_i = \lambda \mathbf{v}_i$, but this time by $\mathbf{v}_j$:

$$\mathbf{v}_j \otimes \mathbf{v}_i + \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \cdot \lambda_j \mathbf{v}_j = 0 + \lambda_i \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \cdot \mathbf{v}_j.$$

$$(\lambda_j - \lambda_i) \frac{\partial \mathbf{v}_i}{\partial \mathbb{A}} \cdot \mathbf{v}_j = -\mathbf{v}_j \otimes \mathbf{v}_i.$$

We also need $(\mathbf{v}_j \otimes \mathbf{v}_i)\mathbb{X}_{ij} = \ldots = \mathbb{P}_i \mathbb{X} \mathbb{P}_j$:

$$\ldots = (\mathbf{v}_j \otimes \mathbf{v}_i)(\mathbf{v}_i \cdot \mathbb{X}\mathbf{v}_j) = (\mathbf{v}_j \otimes \mathbf{v}_i)\mathbb{X}(\mathbf{v}_j \otimes \mathbf{v}_j).$$

□