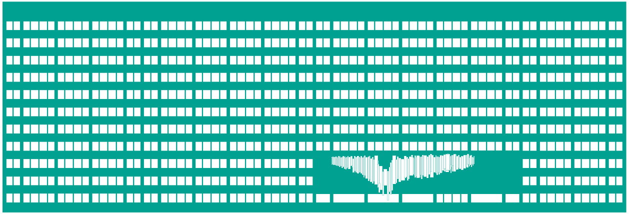
VŠB TECHNICKÁ UNIVERZITA OSTRAVA

VSB TECHNICAL UNIVERSITY OF OSTRAVA



www.vsb.cz

Numerická lineární algebra 1

Aplikace QR rozkladu, mocninná metoda

Michal Merta

VŠB – Technická univerzita Ostrava

michal.merta@vsb.cz

27. dubna 2023

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | IT4INNOVATIONS NÁRODNÍ SUPERPOČÍTAČOVÉ CENTRUM

Metoda nejmenších čtverců – motivační příklad

- Experimentálně jsme zjistili, že tlak v automatickém espresso kávovaru závisí mj. na hrubosti mletí kávových zrn. Závislost můžeme zakreslit do grafu.

Metoda nejmenších čtverců – motivační příklad

- Experimentálně jsme zjistili, že tlak v automatickém espresso kávovaru závisí mj. na hrubosti mletí kávových zrn. Závislost můžeme zakreslit do grafu.
- Graf naznačuje lineární závislost.

Metoda nejmenších čtverců – motivační příklad

- Experimentálně jsme zjistili, že tlak v automatickém espresso kávovaru závisí mj. na hrubosti mletí kávových zrn. Závislost můžeme zakreslit do grafu.
- Graf naznačuje lineární závislost.
- Měření je zatížené chybou – nenalezneme žádnou linární funkci, která prochází všemi body.

- Nedokážeme nalézt funci $p(x) = c_1 + c_2 x$, která prochází všemi body.

- Nedokážeme nalézt funci $p(x) = c_1 + c_2 x$, která prochází všemi body.
- V případě dvou měření $(x_1, y_1), (x_2, y_2)$ bychom získali bychom soustavu dvou rovnic o dvou neznámých

$$\begin{aligned} y_1 &= c_1 + c_2 x_1 \\ y_2 &= c_1 + c_2 x_2, \end{aligned}$$

- Nedokážeme nalézt funci $p(x) = c_1 + c_2 x$, která prochází všemi body.
- V případě dvou měření $(x_1, y_1), (x_2, y_2)$ bychom získali bychom soustavu dvou rovnic o dvou neznámých

$$\begin{aligned} y_1 &= c_1 + c_2 x_1 \\ y_2 &= c_1 + c_2 x_2, \end{aligned}$$

- V našem případě ale získáme pět rovnic o dvou neznámých – přeurčený systém, který obecně nemusí mít řešení.

- Nedokážeme nalézt funci $p(x) = c_1 + c_2 x$, která prochází všemi body.
- V případě dvou měření $(x_1, y_1), (x_2, y_2)$ bychom získali bychom soustavu dvou rovnic o dvou neznámých

$$y_1 = c_1 + c_2 x_1$$
$$y_2 = c_1 + c_2 x_2,$$

- V našem případě ale získáme pět rovnic o dvou neznámých – přeurčený systém, který obecně nemusí mít řešení.
- Můžeme se ale pokusit nalézt přímu p, která sice neprochází všemi body, ale chyba, které se dopouštíme, je v jistém smyslu minimální.

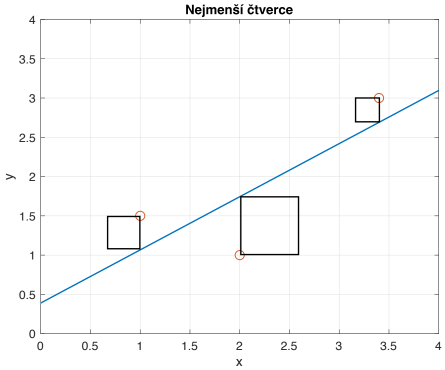
- Hledáme lineární funkci p, která nemusí procházet danými body, ale chyba, které se dopouštíme, je minimální ve smyslu nejmenších čtverců.

- Hledáme lineární funkci p, která nemusí procházet danými body, ale chyba, které se dopouštíme, je minimální ve smyslu nejmenších čtverců.
- Chceme, aby součet druhých mocnin (čtverců) odchylek hledané funkce od naměřených hodnot byl co nejmenší:

$$\min_{p \text{ linearni}} \sum_{j=1}^{m} \left(p(x_j) - y_j\right)^2$$

- Hledáme lineární funkci p, která nemusí procházet danými body, ale chyba, které se dopouštíme, je minimální ve smyslu nejmenších čtverců.
- Chceme, aby součet druhých mocnin (čtverců) odchylek hledané funkce od naměřených hodnot byl co nejmenší:

$$\min_{p \text{ linearni}} \sum_{j=1}^{m} (p(x_j) - y_j)^2$$



- Dosadíme-li $p(x_j) = c_1 + c_2 x_j$ do $\min_{p \text{ linearni}} \sum_{j=1}^{m} (p(x_j) - y_j)^2$, můžeme minimalizační problém přepsat na

$$\min_{c_1, c_2 \in \mathbb{R}} \left\| \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_{m-1} \\ 1 & x_m \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} \right\|^2$$

- Dosadíme-li $p(x_j) = c_1 + c_2 x_j$ do $\min_{p \text{ linearni}} \sum_{j=1}^{m} (p(x_j) - y_j)^2$, můžeme minimalizační problém přepsat na

$$\min_{c_1, c_2 \in \mathbb{R}} \left\| \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_{m-1} \\ 1 & x_m \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} \right\|^2$$

- Problém nejmenších čtverců nemusí vznikat pouze aproximací funkce polynomem – obecně, máme-li $\mathsf{A} \in \mathbb{R}^{m \times n}, \boldsymbol{c} \in \mathbb{R}^n$ a $\boldsymbol{y} \in \mathbb{R}^m, n < m$ a chceme-li aproximovat řešení přeurčeného systému $\mathsf{A}\boldsymbol{c} = \boldsymbol{y}$, získáme problém

$$\min_{\boldsymbol{c} \in \mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|.$$

- Dosadíme-li $p(x_j) = c_1 + c_2 x_j$ do $\min_{p \text{ linearni}} \sum_{j=1}^{m} (p(x_j) - y_j)^2$, můžeme minimalizační problém přepsat na

$$\min_{c_1, c_2 \in \mathbb{R}} \left\| \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ & \vdots \\ 1 & x_{m-1} \\ 1 & x_m \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} \right\|^2$$

- Problém nejmenších čtverců nemusí vznikat pouze aproximací funkce polynomem – obecně, máme-li A $\in \mathbb{R}^{m \times n}, \boldsymbol{c} \in \mathbb{R}^n$ a $\boldsymbol{y} \in \mathbb{R}^m, n < m$ a chceme-li aproximovat řešení přeurčeného systému A$\boldsymbol{c} = \boldsymbol{y}$, získáme problém

$$\min_{\boldsymbol{c} \in \mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|.$$

- Tento problém můžeme řešit pomocí QR rozkladu.

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{Q}\hat{\mathsf{R}}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{Q}\mathsf{R}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.Pak

$$\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\| \quad = \quad \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{Q}\hat{\mathsf{R}}\boldsymbol{c} - \boldsymbol{y}\|$$

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{Q}\mathsf{R}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.Pak

$$\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\| \quad = \quad \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{Q}\hat{\mathsf{R}}\boldsymbol{c} - \boldsymbol{y}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\underbrace{\mathsf{Q}^T\mathsf{Q}}_{=\mathsf{I}}\hat{\mathsf{R}}\boldsymbol{c} - \mathsf{Q}^T\boldsymbol{y}\|$$

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{Q}\hat{\mathsf{R}}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.Pak

$$\begin{aligned} \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\| &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{Q}\hat{\mathsf{R}}\boldsymbol{c} - \boldsymbol{y}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\underbrace{\mathsf{Q}^T\mathsf{Q}}_{=\mathsf{I}}\hat{\mathsf{R}}\boldsymbol{c} - \mathsf{Q}^T\boldsymbol{y}\| \\ &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\hat{\mathsf{R}}\boldsymbol{c} - \underbrace{\mathsf{Q}^T\boldsymbol{y}}_{\overset{\text{ozn.}}{=}\hat{\boldsymbol{y}}}\| \end{aligned}$$

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{QR}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.Pak

$$\begin{aligned} \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\| &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{Q}\hat{\mathsf{R}}\boldsymbol{c} - \boldsymbol{y}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\underbrace{\mathsf{Q}^T\mathsf{Q}}_{=\mathsf{I}}\hat{\mathsf{R}}\boldsymbol{c} - \mathsf{Q}^T\boldsymbol{y}\| \\ &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\hat{\mathsf{R}}\boldsymbol{c} - \underbrace{\mathsf{Q}^T\boldsymbol{y}}_{\overset{\text{ozn.}}{=}\hat{\boldsymbol{y}}}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\hat{\mathsf{R}}\boldsymbol{c} - \hat{\boldsymbol{y}}\| \end{aligned}$$

- Hledejme tedy řešení $\min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\|$ pomocí QR rozkladu.
- Předpokládejme, že známe úplný QR rozklad $\mathsf{A} = \mathsf{Q}\mathsf{R}$, kde

$$\hat{\mathsf{R}} = \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix}$$

a $\mathsf{Q} \in \mathbb{R}^{m\times m}, \mathsf{R} \in \mathbb{R}^{n\times n}, \mathsf{O} \in \mathbb{R}^{(m-n)\times n}$.Pak

$$\begin{aligned} \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{A}\boldsymbol{c} - \boldsymbol{y}\| &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\mathsf{Q}\hat{\mathsf{R}}\boldsymbol{c} - \boldsymbol{y}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\underbrace{\mathsf{Q}^T\mathsf{Q}}_{=\mathsf{I}}\hat{\mathsf{R}}\boldsymbol{c} - \mathsf{Q}^T\boldsymbol{y}\| \\ &= \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\hat{\mathsf{R}}\boldsymbol{c} - \underbrace{\mathsf{Q}^T\boldsymbol{y}}_{\overset{\text{ozn.}}{=}\hat{\boldsymbol{y}}}\| = \min_{\boldsymbol{c}\in\mathbb{R}^m} \|\hat{\mathsf{R}}\boldsymbol{c} - \hat{\boldsymbol{y}}\| \end{aligned}$$

- Náš minimalizační problém jsme tedy převedli na

$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \hat{\boldsymbol{y}}(1:n) \\ \hat{\boldsymbol{y}}(n+1:m) \end{bmatrix} \right\|$$

- Problém

$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \hat{\boldsymbol{y}}(1:n) \\ \hat{\boldsymbol{y}}(n+1:m) \end{bmatrix} \right\|$$

  lze snadno vyřešit.

- Problém
$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \hat{\boldsymbol{y}}(1:n) \\ \hat{\boldsymbol{y}}(n+1:m) \end{bmatrix} \right\|$$
lze snadno vyřešit.
- Prvky na pozicích $n+1:m$ výrazu uvnitř normy žádnou volbou $\boldsymbol{c}$ neovlivníme.

- Problém
$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \hat{\boldsymbol{y}}(1:n) \\ \hat{\boldsymbol{y}}(n+1:m) \end{bmatrix} \right\|$$
lze snadno vyřešit.
- Prvky na pozicích $n+1:m$ výrazu uvnitř normy žádnou volbou $\boldsymbol{c}$ neovlivníme.
- Prvních n prvků můžeme vynulovat položíme-li
$$\mathsf{R}\boldsymbol{c} = \hat{\boldsymbol{y}}(1:n),$$

- Problém

$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathsf{R} \\ \mathsf{O} \end{bmatrix} \boldsymbol{c} - \begin{bmatrix} \hat{\boldsymbol{y}}(1:n) \\ \hat{\boldsymbol{y}}(n+1:m) \end{bmatrix} \right\|$$

  lze snadno vyřešit.
- Prvky na pozicích $n+1:m$ výrazu uvnitř normy žádnou volbou $\boldsymbol{c}$ neovlivníme.
- Prvních n prvků můžeme vynulovat položíme-li

$$\mathsf{R}\boldsymbol{c} = \hat{\boldsymbol{y}}(1:n),$$

- Takto získáme hledané minimum.
- Jedná se o systém s horní trojúhelníkovou maticí – řešení získáme pomocí zpětné substituce.

- Řešeni motivačního problému

Spektrální rozklad

Definice

Buď A $\in \mathbb{R}^{n\times n}$. Nenulový vektor $\boldsymbol{v}$ nazveme vlastním vektorem a λ nazveme příslušným vlastním číslem, platí-li

$$\mathrm{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

Spektrální rozklad

Definice

Buď $\mathsf{A} \in \mathbb{R}^{n \times n}$. Nenulový vektor $\boldsymbol{v}$ nazveme vlastním vektorem a λ nazveme příslušným vlastním číslem, platí-li

$$\mathsf{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

- Rovnost můžeme zapsat maticově

$$\mathsf{AV} = \mathsf{V}\Lambda,$$

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \mathsf{V} = (\boldsymbol{v}_1 \,|\, \boldsymbol{v}_2 \,|\, \cdots \,|\, \boldsymbol{v}_n)$$

- spektrum – množina všech vlastních čísel

- V některých případech má $\mathsf{A} \in \mathbb{R}^{m \times m}$ m lineárně nezávislých vlastních vektorů, které mohou být zvoleny tak, že jsou ortonormální.

- V některých případech má $\mathsf{A} \in \mathbb{R}^{m\times m}$ m lineárně nezávislých vlastních vektorů, které mohou být zvoleny tak, že jsou ortonormální.
- V takovém případě existuje rozklad $\mathsf{A} = \mathsf{QDQ}^T$

- V některých případech má $A \in \mathbb{R}^{m \times m}$ m lineárně nezávislých vlastních vektorů, které mohou být zvoleny tak, že jsou ortonormální.
- V takovém případě existuje rozklad $A = QDQ^T$

Věta

Buď $A \in \mathbb{R}^{m \times m}$ je reálná symetrická matice. Pak existují ortogonální matice Q a diagonální matice D takové, že

$$A = QDQ^T$$

Diagonální prvky D jsou vlastní čísla A, sloupce Q jsou ortonormální vlastní vektory A.

- V některých případech má $A \in \mathbb{R}^{m \times m}$ m lineárně nezávislých vlastních vektorů, které mohou být zvoleny tak, že jsou ortonormální.
- V takovém případě existuje rozklad $A = QDQ^T$

Věta

Buď $A \in \mathbb{R}^{m \times m}$ je reálná symetrická matice. Pak existují ortogonální matice Q a diagonální matice D takové, že

$$A = QDQ^T$$

Diagonální prvky D jsou vlastní čísla A, sloupce Q jsou ortonormální vlastní vektory A.

- Symetrické matice nejsou jediné, pro které existuje $A = QDQ^T$. Obecně platí: A je normální matice (tzn. $A^T A = AA^T$) $\Leftrightarrow$ existuje $A = QDQ^T$ (matice ale mohou být komplexní)

- Spektrální rozklad pomocí QR algoritmu
  - nápad: zkusme použít QR pomocí Householderových transformací pro vynulování nejen prvků ve sloupci, ale i v odpovídajícím řádku

- Spektrální rozklad pomocí QR algoritmu
  - nápad: zkusme použít QR pomocí Householderových transformací pro vynulování nejen prvků ve sloupci, ale i v odpovídajícím řádku

$$\mathsf{A} = \mathsf{Q}_1\mathsf{R}_1, \quad \mathsf{Q}_1^T\mathsf{A} = \mathsf{R}_1 = \begin{pmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}$$

- Spektrální rozklad pomocí QR algoritmu
  - nápad: zkusme použít QR pomocí Householderových transformací pro vynulování nejen prvků ve sloupci, ale i v odpovídajícím řádku

$$\mathsf{A} = \mathsf{Q}_1\mathsf{R}_1, \quad \mathsf{Q}_1^T\mathsf{A} = \mathsf{R}_1 = \begin{pmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}$$

  ale

$$\mathsf{Q}_1^T\mathsf{A}\mathsf{Q}_1 = \mathsf{R}_1\mathsf{Q}_1 = \begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix}$$

  - Násobení zleva opět vytvoří nenulové prvky i v prvním sloupci.

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$A_0 = A$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$
\begin{aligned}
\mathsf{A}_0 &= \mathsf{A} \\
\mathsf{Q}_k \mathsf{R}_k &= \mathsf{A}_{k-1}
\end{aligned}
$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$
\begin{aligned}
\mathsf{A}_0 &= \mathsf{A} \\
\mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\
\mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k
\end{aligned}
$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k \mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k \mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\mathsf{A}_k = \mathsf{R}_k \mathsf{Q}_k$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\mathsf{A}_k = \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{-1}\underbrace{\mathsf{Q}_k\mathsf{R}_k}_{=\mathsf{A}_{k-1}}\mathsf{Q}_k$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} A_0 &= A \\ Q_k R_k &= A_{k-1} \\ A_k &= R_k Q_k \end{aligned}$$

- V k. kroku tedy získáme:

$$A_k = R_k Q_k = Q_k^{-1} \underbrace{Q_k R_k}_{=A_{k-1}} Q_k = Q_k^{-1} A_{k-1} Q_k$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\mathsf{A}_k = \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{-1}\underbrace{\mathsf{Q}_k\mathsf{R}_k}_{=\mathsf{A}_{k-1}}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{A}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{R}_{k-1}\mathsf{Q}_{k-1}\mathsf{Q}_k$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\mathsf{A}_k = \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{-1}\underbrace{\mathsf{Q}_k\mathsf{R}_k}_{=\mathsf{A}_{k-1}}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{A}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{R}_{k-1}\mathsf{Q}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{Q}_{k-1}^{-1}\underbrace{\mathsf{Q}_{k-1}\mathsf{R}_{k-1}}_{=\mathsf{A}_{k-2}}\mathsf{Q}_{k-1}\mathsf{Q}_k$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\begin{aligned} \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{-1}\underbrace{\mathsf{Q}_k\mathsf{R}_k}_{=\mathsf{A}_{k-1}}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{A}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{R}_{k-1}\mathsf{Q}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{Q}_{k-1}^{-1}\underbrace{\mathsf{Q}_{k-1}\mathsf{R}_{k-1}}_{=\mathsf{A}_{k-2}}\mathsf{Q}_{k-1}\mathsf{Q}_k = \\ &= \mathsf{Q}_k^T\mathsf{Q}_{k-1}^T\mathsf{A}_{k-2}\mathsf{Q}_{k-1}\mathsf{Q}_k \end{aligned}$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} \mathsf{A}_0 &= \mathsf{A} \\ \mathsf{Q}_k\mathsf{R}_k &= \mathsf{A}_{k-1} \\ \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\begin{aligned} \mathsf{A}_k &= \mathsf{R}_k\mathsf{Q}_k = \mathsf{Q}_k^{-1}\underbrace{\mathsf{Q}_k\mathsf{R}_k}_{=\mathsf{A}_{k-1}}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{A}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{R}_{k-1}\mathsf{Q}_{k-1}\mathsf{Q}_k = \mathsf{Q}_k^{-1}\mathsf{Q}_{k-1}^{-1}\underbrace{\mathsf{Q}_{k-1}\mathsf{R}_{k-1}}_{=\mathsf{A}_{k-2}}\mathsf{Q}_{k-1}\mathsf{Q}_k = \\ &= \mathsf{Q}_k^T\mathsf{Q}_{k-1}^T\mathsf{A}_{k-2}\mathsf{Q}_{k-1}\mathsf{Q}_k = \underbrace{\mathsf{Q}_k^T\mathsf{Q}_{k-1}^T\cdots\mathsf{Q}_1^T}_{=\tilde{\mathsf{Q}}_k^T}\mathsf{A}\underbrace{\mathsf{Q}_1\cdots\mathsf{Q}_{k-1}\mathsf{Q}_k}_{=\tilde{\mathsf{Q}}_k} \end{aligned}$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$
\begin{aligned}
A_0 &= A \\
Q_k R_k &= A_{k-1} \\
A_k &= R_k Q_k
\end{aligned}
$$

- V k. kroku tedy získáme:

$$
\begin{aligned}
A_k &= R_k Q_k = Q_k^{-1} \underbrace{Q_k R_k}_{=A_{k-1}} Q_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} R_{k-1} Q_{k-1} Q_k = Q_k^{-1} Q_{k-1}^{-1} \underbrace{Q_{k-1} R_{k-1}}_{=A_{k-2}} Q_{k-1} Q_k = \\
&= Q_k^T Q_{k-1}^T A_{k-2} Q_{k-1} Q_k = \underbrace{Q_k^T Q_{k-1}^T \cdots Q_1^T}_{=\tilde{Q}_k^T} A \underbrace{Q_1 \cdots Q_{k-1} Q_k}_{=\tilde{Q}_k}
\end{aligned}
$$

tedy

$$A = \tilde{Q}_k A_k \tilde{Q}_k^T$$

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} A_0 &= A \\ Q_k R_k &= A_{k-1} \\ A_k &= R_k Q_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\begin{aligned} A_k &= R_k Q_k = Q_k^{-1} \underbrace{Q_k R_k}_{=A_{k-1}} Q_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} R_{k-1} Q_{k-1} Q_k = Q_k^{-1} Q_{k-1}^{-1} \underbrace{Q_{k-1} R_{k-1}}_{=A_{k-2}} Q_{k-1} Q_k = \\ &= Q_k^T Q_{k-1}^T A_{k-2} Q_{k-1} Q_k = \underbrace{Q_k^T Q_{k-1}^T \cdots Q_1^T}_{=\tilde{Q}_k^T} A \underbrace{Q_1 \cdots Q_{k-1} Q_k}_{=\tilde{Q}_k} \end{aligned}$$

tedy

$$A = \tilde{Q}_k A_k \tilde{Q}_k^T$$

- Matice A a A_k jsou tedy podobné (takže mají stejná vlastní čísla)

- Lze ukázat, že následující postup (QR algoritmus) konverguje ke spektrálnímu rozkladu

$$\begin{aligned} A_0 &= A \\ Q_k R_k &= A_{k-1} \\ A_k &= R_k Q_k \end{aligned}$$

- V k. kroku tedy získáme:

$$\begin{aligned} A_k &= R_k Q_k = Q_k^{-1} \underbrace{Q_k R_k}_{=A_{k-1}} Q_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} R_{k-1} Q_{k-1} Q_k = Q_k^{-1} Q_{k-1}^{-1} \underbrace{Q_{k-1} R_{k-1}}_{=A_{k-2}} Q_{k-1} Q_k = \\ &= Q_k^T Q_{k-1}^T A_{k-2} Q_{k-1} Q_k = \underbrace{Q_k^T Q_{k-1}^T \cdots Q_1^T}_{=\tilde{Q}_k^T} A \underbrace{Q_1 \cdots Q_{k-1} Q_k}_{=\tilde{Q}_k} \end{aligned}$$

tedy

$$A = \tilde{Q}_k A_k \tilde{Q}_k^T$$

- Matice A a A_k jsou tedy podobné (takže mají stejná vlastní čísla)
- Pro $k \to \infty$: $\tilde{Q}_k \to Q, A_k \to D$ a platí $A = QDQ^T$, kde D je diagonální a Q je ortogonální matice tvořená vlastními vektry A.

function QDQT($A \in \mathbb{R}^{m \times m}$)

```
    D=A
    while norm(D - diag(diag(D))>eps) do
        [Qk, Rk] = qr(D)
        D = Rk*Qk
        Q=Q*Qk
    end while
end function
```

function QDQT($A \in \mathbb{R}^{m\times m}$)
```
    D=A
    while norm(D - diag(diag(D))>eps) do
        [Qk, Rk] = qr(D)
        D = Rk*Qk
        Q=Q*Qk
    end while
end function
```

- Pozn. V případě nesymetrické matice konverguje A_k k trojúhelníkové matici s vlastními čísly na diagonále.

Mocninná metoda

- Řešme tzv. částečný problém vlastních čísel (hledejme největší vlastní číslo v absolutní hodnotě).

Definice

Buď $\lambda_1, \lambda_2, \ldots, \lambda_n$ vlastní čísla čísla matice A $\in \mathbb{R}^{n\times n}$. Pak λ_1 nazveme dominantním vlastním číslem A, platí-li

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|.$$

Vlastní vektor odpovídající dominantnímu vlastnímu číslu se nazývá dominantní vlastní vektor.

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.
  3. Vytvoříme posloupnost aproximací ve tvaru

$$\boldsymbol{x}^1 = \mathsf{A}\boldsymbol{x}^0$$

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.
  3. Vytvoříme posloupnost aproximací ve tvaru

$$\begin{aligned}\boldsymbol{x}^1 &= \mathsf{A}\boldsymbol{x}^0\\ \boldsymbol{x}^2 &= \mathsf{A}\boldsymbol{x}^1 = \mathsf{A}(\mathsf{A}\boldsymbol{x}^0) = \mathsf{A}^2\boldsymbol{x}^0\end{aligned}$$

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.
  3. Vytvoříme posloupnost aproximací ve tvaru

$$
\begin{aligned}
\boldsymbol{x}^1 &= \mathsf{A}\boldsymbol{x}^0 \\
\boldsymbol{x}^2 &= \mathsf{A}\boldsymbol{x}^1 = \mathsf{A}(\mathsf{A}\boldsymbol{x}^0) = \mathsf{A}^2\boldsymbol{x}^0 \\
\boldsymbol{x}^3 &= \mathsf{A}\boldsymbol{x}^2 = \mathsf{A}(\mathsf{A}(\mathsf{A}\boldsymbol{x}^0)) = \mathsf{A}^3\boldsymbol{x}^0
\end{aligned}
$$

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.
  3. Vytvoříme posloupnost aproximací ve tvaru

$$
\begin{aligned}
\boldsymbol{x}^1 &= \mathsf{A}\boldsymbol{x}^0 \\
\boldsymbol{x}^2 &= \mathsf{A}\boldsymbol{x}^1 = \mathsf{A}(\mathsf{A}\boldsymbol{x}^0) = \mathsf{A}^2\boldsymbol{x}^0 \\
\boldsymbol{x}^3 &= \mathsf{A}\boldsymbol{x}^2 = \mathsf{A}(\mathsf{A}(\mathsf{A}\boldsymbol{x}^0)) = \mathsf{A}^3\boldsymbol{x}^0 \\
&\vdots \\
\boldsymbol{x}^k &= \mathsf{A}\boldsymbol{x}^{k-1} = \mathsf{A}(\mathsf{A}^{k-1}\boldsymbol{x}^0) = \mathsf{A}^k\boldsymbol{x}^0
\end{aligned}
$$

- K nalezení dominantního vlastního čísla a vektoru můžeme použít mocninnou metodu. Její princip je jednoduchý:
  1. Předpokládáme, že A má dominantní vlastní číslo.
  2. Zvolíme nenulovou počáteční aproximaci $\boldsymbol{x}^0$ dominantního vlastního vektoru.
  3. Vytvoříme posloupnost aproximací ve tvaru

$$\begin{aligned}
\boldsymbol{x}^1 &= \mathsf{A}\boldsymbol{x}^0 \\
\boldsymbol{x}^2 &= \mathsf{A}\boldsymbol{x}^1 = \mathsf{A}(\mathsf{A}\boldsymbol{x}^0) = \mathsf{A}^2\boldsymbol{x}^0 \\
\boldsymbol{x}^3 &= \mathsf{A}\boldsymbol{x}^2 = \mathsf{A}(\mathsf{A}(\mathsf{A}\boldsymbol{x}^0)) = \mathsf{A}^3\boldsymbol{x}^0 \\
&\vdots \\
\boldsymbol{x}^k &= \mathsf{A}\boldsymbol{x}^{k-1} = \mathsf{A}(\mathsf{A}^{k-1}\boldsymbol{x}^0) = \mathsf{A}^k\boldsymbol{x}^0
\end{aligned}$$

- Tato posloupnost při splnění určitých podmínek konverguje k dominantnímu vlastnímu vektoru.

- Omezme se na reálné symetrické matice A. Ty mají n lineárně nezávislých vlastních vektorů a reálná vlastní čísla.

Vět

Je-li $\boldsymbol{x}$ vlastní vektor reálné symetrické matice A, pak odpovídající vlastní číslo je dáno výrazem

$$\lambda = \frac{(\mathsf{A}\boldsymbol{x})^T\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}.$$

- Omezme se na reálné symetrické matice A. Ty mají n lineárně nezávislých vlastních vektorů a reálná vlastní čísla.

Vět

Je-li $\boldsymbol{x}$ vlastní vektor reálné symetrické matice A, pak odpovídající vlastní číslo je dáno výrazem

$$\lambda = \frac{(\mathrm{A}\boldsymbol{x})^T \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

Důkaz

$$\frac{(\mathrm{A}\boldsymbol{x})^T \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{(\lambda\boldsymbol{x})^T \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \lambda$$

$\square$

Věta

Buď $\mathsf{A} \in \mathbb{R}^{n\times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0 \mathsf{A}^3\boldsymbol{x}^0, \ldots, \mathsf{A}^k\boldsymbol{x}^0, \ldots$$

konverguje k násobku dominantního vlastního vektoru A.

Věta

Buď $\mathsf{A} \in \mathbb{R}^{n \times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0 \mathsf{A}^3\boldsymbol{x}^0, \ldots, \mathsf{A}^k\boldsymbol{x}^0, \ldots$$

konverguje k násobku dominantního vlastního vektoru A.

Důkaz.

- Předpokládejme, že pro vlastní čísla matice A platí $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|$. Odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n$ jsou lineárně nezávislé a tvoří bázi $\mathbb{R}^n$.

Věta

Buď A $\in \mathbb{R}^{n\times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0\mathsf{A}^3\boldsymbol{x}^0, \ldots, \mathsf{A}^k\boldsymbol{x}^0, \ldots$$

konverguje k násobku dominantního vlastního vektoru A.

Důkaz.

- Předpokládejme, že pro vlastní čísla matice A platí $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|$. Odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n$ jsou lineárně nezávislé a tvoří bázi $\mathbb{R}^n$.
- Zvolme vektor $\boldsymbol{x}^0$ tak, aby koeficient c_1 v lineární kombinaci

$$\boldsymbol{x}^0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$$

  byl nenulový (pro $c_1 = 0$ nemusí metoda konvergovat).

Věta

Buď $\mathsf{A} \in \mathbb{R}^{n \times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0 \mathsf{A}^3\boldsymbol{x}^0, \ldots, \mathsf{A}^k\boldsymbol{x}^0, \ldots$$

konverguje k násobku dominantního vlastního vektoru A.

Důkaz.

- Předpokládejme, že pro vlastní čísla matice A platí $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|$. Odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n$ jsou lineárně nezávislé a tvoří bázi $\mathbb{R}^n$.
- Zvolme vektor $\boldsymbol{x}^0$ tak, aby koeficient c_1 v lineární kombinaci

  $$\boldsymbol{x}^0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$$

  byl nenulový (pro $c_1 = 0$ nemusí metoda konvergovat).
- Přenásobme předchozí rovnost maticí A

  $$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n)$$

Věta

Buď $\mathsf{A} \in \mathbb{R}^{n\times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0 \mathsf{A}^3\boldsymbol{x}^0, \dots, \mathsf{A}^k\boldsymbol{x}^0, \dots$$

konverguje k násobku dominantního vlastního vektoru A.

Důkaz.

- Předpokládejme, že pro vlastní čísla matice A platí $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots |\lambda_n|$. Odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_n$ jsou lineárně nezávislé a tvoří bázi $\mathbb{R}^n$.
- Zvolme vektor $\boldsymbol{x}^0$ tak, aby koeficient c_1 v lineární kombinaci

$$\boldsymbol{x}^0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \dots + c_n\boldsymbol{v}_n$$

  byl nenulový (pro $c_1 = 0$ nemusí metoda konvergovat).
- Přenásobme předchozí rovnost maticí A

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \dots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \dots + c_n\mathsf{A}\boldsymbol{v}_n$$

Věta

Buď $\mathsf{A} \in \mathbb{R}^{n \times n}$ matice s n lineárně nezávislými vlastními vektory a dominantním vlastním číslem. Pak existuje nenulový vektor $\boldsymbol{x}^0$ takový, že posloupnost

$$\mathsf{A}\boldsymbol{x}^0, \mathsf{A}^2\boldsymbol{x}^0 \mathsf{A}^3\boldsymbol{x}^0, \ldots, \mathsf{A}^k\boldsymbol{x}^0, \ldots$$

konverguje k násobku dominantního vlastního vektoru A.

Důkaz.

- Předpokládejme, že pro vlastní čísla matice A platí $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|$. Odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n$ jsou lineárně nezávislé a tvoří bázi $\mathbb{R}^n$.
- Zvolme vektor $\boldsymbol{x}^0$ tak, aby koeficient c_1 v lineární kombinaci
$$\boldsymbol{x}^0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$$
byl nenulový (pro $c_1 = 0$ nemusí metoda konvergovat).
- Přenásobme předchozí rovnost maticí A
$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\Downarrow$$

$$\mathsf{A}^k\boldsymbol{x}^0 \quad = \quad c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \ldots + c_n\lambda_n^k\boldsymbol{v}_n$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\Downarrow$$

$$\begin{aligned}\mathsf{A}^k\boldsymbol{x}^0 &= c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \ldots + c_n\lambda_n^k\boldsymbol{v}_n \\ &= \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\frac{\lambda_2^k}{\lambda_1^k}\boldsymbol{v}_2 + \ldots + c_n\frac{\lambda_n^k}{\lambda_1^k}\boldsymbol{v}_n\right)\end{aligned}$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\Downarrow$$

$$\begin{aligned} \mathsf{A}^k\boldsymbol{x}^0 &= c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \ldots + c_n\lambda_n^k\boldsymbol{v}_n \\ &= \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\frac{\lambda_2^k}{\lambda_1^k}\boldsymbol{v}_2 + \ldots + c_n\frac{\lambda_n^k}{\lambda_1^k}\boldsymbol{v}_n\right) = \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2 + \ldots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n\right). \end{aligned}$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\Downarrow$$

$$\begin{aligned} \mathsf{A}^k\boldsymbol{x}^0 &= c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \ldots + c_n\lambda_n^k\boldsymbol{v}_n \\ &= \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\frac{\lambda_2^k}{\lambda_1^k}\boldsymbol{v}_2 + \ldots + c_n\frac{\lambda_n^k}{\lambda_1^k}\boldsymbol{v}_n\right) = \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2 + \ldots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n\right). \end{aligned}$$

- Protože $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|.$, platí

$$\left|\frac{\lambda_2}{\lambda_1}\right| < 1, \ldots, \left|\frac{\lambda_n}{\lambda_1}\right| < 1.$$

$$\mathsf{A}\boldsymbol{x}^0 = \mathsf{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) = c_1\mathsf{A}\boldsymbol{v}_1 + c_2\mathsf{A}\boldsymbol{v}_2 + \ldots + c_n\mathsf{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_n\lambda_n\boldsymbol{v}_n.$$

$$\Downarrow$$

$$\begin{aligned} \mathsf{A}^k\boldsymbol{x}^0 &= c_1\lambda_1^k\boldsymbol{v}_1 + c_2\lambda_2^k\boldsymbol{v}_2 + \ldots + c_n\lambda_n^k\boldsymbol{v}_n \\ &= \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\frac{\lambda_2^k}{\lambda_1^k}\boldsymbol{v}_2 + \ldots + c_n\frac{\lambda_n^k}{\lambda_1^k}\boldsymbol{v}_n\right) = \lambda_1^k\left(c_1\boldsymbol{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2 + \ldots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n\right). \end{aligned}$$

- Protože $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots |\lambda_n|.$, platí

$$\left|\frac{\lambda_2}{\lambda_1}\right| < 1, \ldots, \left|\frac{\lambda_n}{\lambda_1}\right| < 1.$$

- Pro $k \to \infty$ tedy výraz $\mathsf{A}^0\boldsymbol{x}^0$ konverguje k $\lambda_1^k c_1\boldsymbol{v}_1$, což je násobek dominantního vlastního vektoru. $\square$

function Mocninna_metoda(A, $\boldsymbol{x}^0$)

  $\boldsymbol{q}^0 = \frac{\boldsymbol{x}^0}{\|\boldsymbol{x}^0\|}$

  for $k = 1, 2, \ldots$ **do**

    $\boldsymbol{x}^k = \mathsf{A}\boldsymbol{q}^{k-1}$

    $\boldsymbol{q}^k = \frac{\boldsymbol{x}^k}{\|\boldsymbol{x}^k\|}$

    $\lambda = (\boldsymbol{q}^k)^T \mathsf{A}\boldsymbol{q}^k \qquad \left(= \frac{(\boldsymbol{x}^k)^T \mathsf{A}\boldsymbol{x}^k}{\|\boldsymbol{x}^k\|^2} = \frac{(\boldsymbol{x}^k)^T \mathsf{A}\boldsymbol{x}^k}{(\boldsymbol{x}^k)^T \boldsymbol{x}^k}\right)$

  end for

end function

function Mocninna_metoda(A, $\boldsymbol{x}^0$)

$\boldsymbol{q}^0 = \frac{\boldsymbol{x}^0}{\|\boldsymbol{x}^0\|}$

for $k = 1, 2, \ldots$ **do**

$\boldsymbol{x}^k = \mathsf{A}\boldsymbol{q}^{k-1}$

$\boldsymbol{q}^k = \frac{\boldsymbol{x}^k}{\|\boldsymbol{x}^k\|}$

$\lambda = (\boldsymbol{q}^k)^T \mathsf{A}\boldsymbol{q}^k \qquad \left(= \frac{(\boldsymbol{x}^k)^T \mathsf{A}\boldsymbol{x}^k}{\|\boldsymbol{x}^k\|^2} = \frac{(\boldsymbol{x}^k)^T \mathsf{A}\boldsymbol{x}^k}{(\boldsymbol{x}^k)^T \boldsymbol{x}^k}\right)$

end for

end function

- Abychom zachovali numerickou stabilitu, vektor v každé iteraci normujeme.
- Jednou z možností, jak zvolit ukončovací podmínku, je sledovat normu vektoru $\mathsf{A}\boldsymbol{q}^k - \lambda\boldsymbol{q}^k$. Ta by se měla blížit nule pro $k \to \infty$.

Posun spektra

- Má-li A vlastní čísla $\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ a odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, jak vypadají vlastní čísla a vektory matice $\mathsf{A} - \sigma \mathsf{I}$, kde $\sigma \in \mathbb{R}$?

Posun spektra

- Má-li A vlastní čísla $\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ a odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, jak vypadají vlastní čísla a vektory matice $\mathsf{A} - \sigma \mathsf{I}$, kde $\sigma \in \mathbb{R}$?
- Vyjděme z rovnosti

$$\mathsf{A}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k$$

- Odečtěme od obou stran σ násobek vlastního vektoru $\boldsymbol{v}_k$:

$$\mathsf{A}\boldsymbol{v}_k - \sigma \mathsf{I}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k - \sigma \boldsymbol{v}_k.$$

Posun spektra

- Má-li A vlastní čísla $\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ a odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, jak vypadají vlastní čísla a vektory matice $\mathsf{A} - \sigma \mathsf{I}$, kde $\sigma \in \mathbb{R}$?
- Vyjděme z rovnosti

$$\mathsf{A}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k$$

- Odečtěme od obou stran σ násobek vlastního vektoru $\boldsymbol{v}_k$:

$$\mathsf{A}\boldsymbol{v}_k - \sigma \mathsf{I}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k - \sigma \boldsymbol{v}_k.$$

$$\Downarrow$$

$$(\mathsf{A} - \sigma \mathsf{I})\boldsymbol{v}_k = (\lambda_k - \sigma)\boldsymbol{v}_k.$$

Posun spektra

- Má-li A vlastní čísla $\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ a odpovídající vlastní vektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, jak vypadají vlastní čísla a vektory matice $\mathsf{A} - \sigma\mathsf{I}$, kde $\sigma \in \mathbb{R}$?
- Vyjděme z rovnosti

$$\mathsf{A}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k$$

- Odečtěme od obou stran σ násobek vlastního vektoru $\boldsymbol{v}_k$:

$$\mathsf{A}\boldsymbol{v}_k - \sigma\mathsf{I}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k - \sigma\boldsymbol{v}_k.$$

$$\Downarrow$$

$$(\mathsf{A} - \sigma\mathsf{I})\boldsymbol{v}_k = (\lambda_k - \sigma)\boldsymbol{v}_k.$$

- $\boldsymbol{v}_k$ je také vlastním vektorem matice $\mathsf{A} - \sigma\mathsf{I}$. Odpovídajícím vlastním číslem je $\lambda_k - \sigma$.

- Předchozího poznatku lze využít k nalezení nejmenšího vlastního čísla λ_n symetrické pozitivně definitní matice A, známe-li její největší vlastní číslo λ_1.

- Předchozího poznatku lze využít k nalezení nejmenšího vlastního čísla λ_n symetrické pozitivně definitní matice A, známe-li její největší vlastní číslo λ_1.
- Sestavme $\mathsf{B} = \mathsf{A} - \lambda_1 \mathsf{I}$.

- Předchozího poznatku lze využít k nalezení nejmenšího vlastního čísla λ_n symetrické pozitivně definitní matice A, známe-li její největší vlastní číslo λ_1.
- Sestavme $\mathsf{B} = \mathsf{A} - \lambda_1 \mathsf{I}$.
- Matice B má dle předchozího poznatku vlastní čísla $\lambda_n - \lambda_1, \lambda_{n-1} - \lambda_1, \ldots$.

- Předchozího poznatku lze využít k nalezení nejmenšího vlastního čísla λ_n symetrické pozitivně definitní matice A, známe-li její největší vlastní číslo λ_1.
- Sestavme $\mathsf{B} = \mathsf{A} - \lambda_1 \mathsf{I}$.
- Matice B má dle předchozího poznatku vlastní čísla $\lambda_n - \lambda_1, \lambda_{n-1} - \lambda_1, \ldots$.
- Předpokládejme, že $\lambda_n \neq \lambda_{n-1}$. Protože symetrická pozitivně definitní matice má všechna vlastní čísla kladná (tedy $0 < \lambda_n < \lambda_{n-1} \leq \ldots$), můžeme vlastní čísla matice s takto posunutým spektrem seřadit

$$|\lambda_n - \lambda_1| > |\lambda_{n-1} - \lambda_1| \geq \ldots .$$

- Předchozího poznatku lze využít k nalezení nejmenšího vlastního čísla λ_n symetrické pozitivně definitní matice A, známe-li její největší vlastní číslo λ_1.
- Sestavme $\mathsf{B} = \mathsf{A} - \lambda_1 \mathsf{I}$.
- Matice B má dle předchozího poznatku vlastní čísla $\lambda_n - \lambda_1, \lambda_{n-1} - \lambda_1, \ldots$.
- Předpokládejme, že $\lambda_n \neq \lambda_{n-1}$. Protože symetrická pozitivně definitní matice má všechna vlastní čísla kladná (tedy $0 < \lambda_n < \lambda_{n-1} \leq \ldots$), můžeme vlastní čísla matice s takto posunutým spektrem seřadit

$$|\lambda_n - \lambda_1| > |\lambda_{n-1} - \lambda_1| \geq \ldots .$$

- Použijeme-li tedy mocninnou metodu na matici $\mathsf{B} = \mathsf{A} - \lambda_1 \mathsf{I}$, nalezneme její dominantní vlastní číslo $\hat{\lambda} = \lambda_n - \lambda_1$. Snadno dopočítáme nejmenší vlastní číslo původní matice A:

$$\lambda_n = \hat{\lambda} + \lambda_1.$$

Děkuji za pozornost

Michal Merta

VŠB – Technická univerzita Ostrava

michal.merta@vsb.cz

27. dubna 2023

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | IT4INNOVATIONS NÁRODNÍ SUPERPOČÍTAČOVÉ CENTRUM