Přednáška 9b: Aplikace pseudoinverze, další rozklady matic, lineární procesy

Minule jsme definovali pseudoinverzní matici k matici A se sing. rozkladem $A = PSQ^*$ jako

$$A^{(-1)} = Q\, S^{(-1)} P^*$$

a uvedli jsme něco o vlastnostech pseudoinverzní matice. Některé z těchto vlastností dokážeme:

(4) $A^{(-1)}A$ je matice kolmé projekce $\mathbb{K}^n$ na $(\ker\varphi)^\perp$

Z důkazu sing. rozkladu víme, že

$Q = (id)_{\varepsilon_n,\alpha}$, kde $\alpha = (u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_n)$

je ortonormální báze v $\mathbb{K}^n$ a $u_{r+1}, \dots, u_n \in \ker\varphi$,

$\varphi : \mathbb{K}^n \to \mathbb{K}^k$, $\varphi(x) = Ax$.

Platí $\qquad (\varphi^{(-1)}\varphi)_{\varepsilon_n,\varepsilon_n} = A^{(-1)}A$

a také

$$(\varphi^{(-1)}\varphi)_{\alpha,\alpha} = (id)_{\alpha,\varepsilon_n} A^{(-1)} A\, (id)_{\varepsilon_n,\alpha}$$

$$= Q^*\left(Q S^{(-1)} P^* P S Q^*\right) Q$$

$$= S^{(-1)} S = \begin{array}{cc} \left(\begin{array}{c|c} E & 0 \\ \hline 0 & 0 \end{array}\right) & \begin{array}{l} \}\, r \\ \}\, n-r \end{array} \\ \begin{array}{cc} \underbrace{\quad}_{r} & \underbrace{\quad}_{n-r} \end{array} & \end{array}$$

Proto $\quad (\varphi^{(-1)}\varphi)(u_i) = u_i \quad$ pro $\quad 1 \le i \le r$

$\qquad\quad (\varphi^{(-1)}\varphi)(u_j) = 0 \quad$ pro $\quad r+1 \le j \le n$

Tedy $\varphi^{(-1)}\varphi$ je kolmá projekce na $(\ker\varphi)^\perp$.

(5) $AA^{(-1)}$ je matice kolmé projekce $\mathbb{K}^k$ na $\operatorname{im}\varphi$.

Připomeňme, že $P = (id)_{\mathcal{E}_k, B}$ kde $B = (v_1, \dots, v_r, v_{r+1}, \dots, v_k)$ je ortonormální báze, kde

$$[v_1, v_2, \dots, v_r] = \operatorname{im}\varphi \,, \quad \varphi(x) = Ax .$$

Platí

$$(\varphi\,\varphi^{(-1)})_{\mathcal{E}_k, \mathcal{E}_k} = A A^{(-1)}$$

$$(\varphi\,\varphi^{(-1)})_{B, B} = (id)_{B, \mathcal{E}_k} \, A A^{(-1)} \, (id)_{\mathcal{E}_k, B} =$$

$$= P^* (P S Q^*)(Q S^{-1} P^*) P =$$

$$= S S^{-1} = \begin{array}{c} \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \begin{array}{l} \}\, r \\ \}\, k-r \end{array} \\ \quad\;\; r \quad\; k-r \qquad\quad \end{array}$$

Proto $\varphi\,\varphi^{(-1)}(v_i) = v_i$ pro $1 \le i \le r$

a

$$\varphi\,\varphi^{(-1)}(v_j) = 0 \quad \text{pro } r+1 \le j \le n .$$

Tedy $\varphi\,\varphi^{(-1)}$ je kolmá projekce na $\operatorname{im}\varphi$.

(6) Platí $A A^{(-1)} A = A$.

$$A A^{(-1)} A = P S Q^* Q S^{(-1)} P^* P S Q^* \qquad S$$

$$= P S S^{(-1)} S Q^* = P \underbrace{\begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}}_{k \times k} \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{k \times n} Q^*$$

$$= P \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{k \times n} Q^* = P S Q^* = A$$

(7) $$A^{(-1)} = (A^*A)^{(-1)}A^*$$

$$(A^*A)^{(-1)}A^* = \left((PSQ^*)^*(PSQ^*)\right)^{(-1)}(PSQ^*)^*$$

$$= (QS^*P^*PSQ^*)^{-1}(QS^*P^*) =$$

$$= \left(Q \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{n\times k} \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{k\times n} Q^*\right)^{(-1)} \cdot Q \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{n\times k} P^* =$$

$$= \left(Q \underbrace{\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix}}_{n\times n} Q^*\right)^{(-1)} \cdot Q \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{n\times k} P^* =$$

$$= Q \underbrace{\begin{pmatrix} D^{-2} & 0 \\ 0 & 0 \end{pmatrix}}_{n\times n} \underbrace{Q^* \quad Q}_{\substack{=E \\ n\times n}} \underbrace{\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}}_{n\times k} P^* =$$

$$= Q \underbrace{\begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}}_{n\times k} P^* = QS^{(-1)}P^* = A^{(-1)}$$

Platí rovněž $A^*(AA^*)^{(-1)} = A^{(-1)}$. Důkaz je domácí úloha.

Aproximace řešení soustav lineárních rovnic

Nechť A je matice $k \times n$. Uvažujme soustavu

$$A x = b, \quad x \in \mathbb{K}^n, \quad b \in \mathbb{K}^k.$$

Tato soustava má řešení, právě když

$$(*) \qquad h(A) = h(A|b)$$

Podmínka $(*)$ je ekvivalentní tomu, že

$$b \in \operatorname{im} \varphi, \quad \varphi : \mathbb{K}^n \to \mathbb{K}^k, \ \varphi(x) = Ax$$

V případě, že podmínka $(*)$ není splněna, chceme najít $x \in \mathbb{K}^n$ takové, že

$$\|Ax - b\| \text{ je minimální.}$$

Věta: Funkce $f : \mathbb{K}^n \to \mathbb{R}$

$$f(x) = \|Ax - b\|$$

nabývá svého minima v bodech

$$x = A^{(-1)} b + y,$$

kde $Ay = 0$.

Důkaz: Již víme, že $A A^{(-1)} = P$ je matice

kolmé projekce $\mathbb{K}^k$ na $\operatorname{im}\varphi$ $(\varphi(x) = Ax)$.
Protože

$$\min_{x \in \mathbb{K}^n} \|Ax - b\| = \min_{v \in \operatorname{im}\varphi} \|v - b\| = \|Pb - b\|$$
$$= \|AA^{(-1)}b - b\|$$

Tedy v $x = A^{(-1)}b$ nabývá funkce f svého minima. Stejné hodnoty ale nabývá i ve všech $A^{(-1)}b + y$, kde $Ay = 0$.

Příklad: Máme (přeurčenou) soustavu rovnic

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ x_1 \phantom{{}+2x_2} &= 7 \\ x_2 &= 5 \end{aligned}$$

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix} \notin \operatorname{im}\varphi$$

Hledáme (x_1, x_2) tak, aby

$$\|Ax - b\|^2 = (x_1 + 2x_2 - 3)^2 + (x_1 - 7)^2 + (x_2 - 5)^2$$

byla minimální. Podle předchozí věty, je

$$x = A^{(-1)}b = \begin{pmatrix} 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 14/3 \\ 1/3 \end{pmatrix}$$

$A^{(-1)}$ jsme spočítali minule!

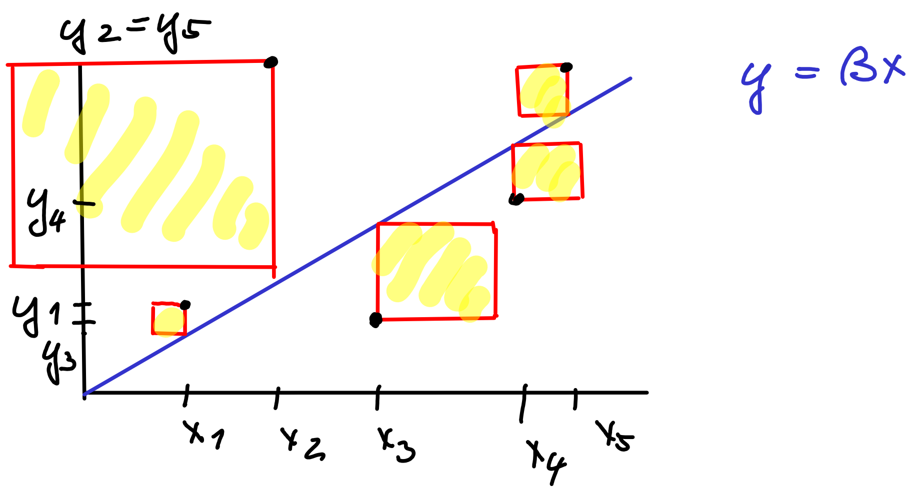
Lineární regrese:

Očekáváme, že závislost veličin x, y je lineární:

$$y = \beta x$$

Naměříme n hodnot $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ a chceme najít β tak, že přímka $y = \beta x$ je „blízko" naměřených bodů (x_i, y_i). Slovo blízko znamená obvykle, že suma

$$\sum_{i=1}^{n} (y_i - \beta x_i)^2 \quad \text{je minimální.}$$



Máme soustavu

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \beta = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

o neznámé $\beta \in \mathbb{R}$.

Nechť $\sum_{i=1}^{n} x_i^2 \neq 0$.
$$A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad b = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\|A\beta - b\|^2 = (\beta x_1 - y_1)^2 + (\beta x_2 - y_2)^2 + \dots + (\beta x_n - y_n)^2 .$$

Nejlepší aproximace je

$$\beta = A^{(-1)} b = (A^* A)^{(-1)} (A^* b)$$

$$= \left(\sum_{i=1}^{n} x_i^2 \right)^{-1} (x_1\, x_2 \dots x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \frac{\sum_{i=1}^{n} x_i \cdot y_i}{\sum_{i=1}^{n} x_i^2} .$$

Další příklad na lineární regresi

Předpokládáme závislost $y = \alpha + \beta x$, naměřili jsme $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

Hledáme α, β tak, aby součet čtverců

$$\sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \quad \text{byl minimální.}$$

To vede k soustavě

$$\begin{aligned} \alpha + \beta x_1 &= y_1 \\ \alpha + \beta x_2 &= y_2 \\ &\dots \\ \alpha + \beta x_n &= y_n \end{aligned}$$

s maticí $A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}$, $b = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, neznámými $\underline{\alpha}$ a $\underline{\beta}$.

Víme, že
$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = A^{(-1)}\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$$

$$A^*A = \begin{pmatrix}1 & 1 & \dots & 1\\ x_1 & x_2 & & x_n\end{pmatrix}\begin{pmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \\ 1 & x_n\end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} x_i\\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2\end{pmatrix}$$

$$\det A^*A = n\sum_{i=1}^{n} x_i - \left(\sum_i x_i\right)^2 = \sum_{1\le i<j\le n} (x_i - x_j)^2$$

Nechť $x_i \neq x_j$ pro nějaké $i \neq j$. Pak $\det A^*A \neq 0$ a proto je

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = (A^*A)^{-1}(A^*b) = \frac{1}{\sum_{i<j}(x_i-x_j)^2}\begin{pmatrix}\sum x_i^2 & \sum x_i\\ \sum x_i & n\end{pmatrix}\begin{pmatrix}\sum_{i=1}^{n} y_i\\ \sum_{i=1}^{n} x_i y_i\end{pmatrix}$$

$$= \frac{1}{\sum_{i<j}(x_i-x_j)^2}\begin{pmatrix}\left(\sum_i x_i^2\right)\left(\sum_j y_j\right) + \left(\sum_i x_i\right)\left(\sum_j x_j y_j\right)\\ \left(\sum_i x_i\right)\left(\sum_j y_j\right) + n\sum_{i=1}^{n} x_i y_i\end{pmatrix}$$

Polární rozklad matice

Motivace: Každé komplexní číslo $a+ib \in \mathbb{C}$ lze psát ve tvaru
$$a + ib = r(\cos\alpha + i\sin\alpha),$$

kde $r \geq 0$. Tento rozklad je jednoznačný, je-li $a+ib \neq 0$.

Uvažujme lineární zobrazení

$$\varphi : \mathbb{C} \longrightarrow \mathbb{C} \qquad \varphi(z) = (a+ib)\cdot z$$

Toto zobrazení je složením dvou zobrazení

$$\varphi = \varphi_2 \circ \varphi_1 ,$$

kde

$$\varphi_1 : \mathbb{C} \longrightarrow \mathbb{C} \qquad \varphi_1(z) = (\cos\alpha + i \sin\alpha)\, z$$

je unitární a

$$\varphi_2 : \mathbb{C} \longrightarrow \mathbb{C}, \qquad \varphi_2(z) = r z \quad , r \geq 0$$

je samoadjungované a pozitivně semidefinitní.

Samoadjungovanost $(r)^* = \overline{r} = r$.

Pozitivní semidefinitnost

$$\langle rz, z \rangle = r\,|z|^2 \geq 0 .$$

Věta o polárním rozkladu

Je-li A čtvercová matice $n \times n$ nad $\mathbb{R}$ nebo $\mathbb{C}$, pak ji lze rozložit na součin

$$A = R \cdot U ,$$

kde R je samoadjungovaná $(R = R^*)$ a pozitivně semidefinitní $(\langle Rx, x \rangle \geq 0)$ a U je unitární nebo ortogonální.

Navíc platí, že $R^2 = AA^*$ (píšeme $R = \sqrt{AA^*}$).
Je-li A invertibilní, jsou R a U určeny jednoznačně.

Důkaz pomocí sing. rozkladu: Nechť $A = PSQ^*$ je sing. rozklad matice A. Pak

$$A = PS\underbrace{P^*P}_{E}Q^* = \underbrace{(PSP^*)}_{R}\underbrace{(PQ^*)}_{U}$$

R je samoadjungovaná a pozitivně semidefinitní:

$$R^* = (PSP^*)^* = PS^*P^* = PSP^*$$

$$\langle Rx, x\rangle = \langle PSP^*x, x\rangle = \langle SP^*x, P^*x\rangle = \langle Sy, y\rangle \geq 0.$$

U je unitární neboli ortogonální:

$$UU^* = (PQ^*)(PQ^*)^* = (PQ^*)(QP^*) = P\underbrace{Q^*Q}_{E}P = E$$

Dále:

$$AA^* = (RU)(RU)^* = R\underbrace{UU^*}_{E}R^* = R \cdot R^* = R^2.$$

Jednoznačnost pro A invertibilní nebudeme dokazovat.

Příklad: Najděte polární rozklad matice

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$$

A není regulární. $\quad A = PSQ^* = P\begin{pmatrix} \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 \end{pmatrix}Q^*$

$$A = (PSP^*)(PQ^*) = \begin{pmatrix} \frac{\sqrt{10}}{2} & \frac{\sqrt{10}}{2} \\ \frac{\sqrt{10}}{2} & \frac{\sqrt{10}}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}$$

Lineární procesy

popisují vývoj nějakého systému, který je v čase n popsán vektorem

$$x(n) = \begin{pmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_m(n) \end{pmatrix} \in \mathbb{R}^m$$

a jeho přechod z času n do času $n+1$ pomocí matice A typu $m \times m$:

$$x(n+1) = A\,x(n) .$$

Důležitou roli pro „vývoj" tohoto systému mají vlastní čísla a vlastní vektory matice A.

1 Dravec a kořist

Uvažujme dva živočišné druhy, dravce D a kořist K. Označme $X(n) = \begin{pmatrix} D_n \\ K_n \end{pmatrix}$ jejich počty po n letech. Změna počtu probíhá takto:

$$\begin{aligned} D_{n+1} &= 0{,}6\,D_n + 0{,}5\,K_n \\ K_{n+1} &= -p\,D_n + 1{,}2\,K_n \end{aligned}$$

kde $p > 0$ je parametr. Maticově

$$X(n+1) = \begin{pmatrix} 0{,}6 & 0{,}5 \\ -p & 1{,}2 \end{pmatrix} X(n) = A\,X(n)$$

Zajímá nás vývoj této dvoudruhové populace v závislosti na parametru p.

Vlastní čísla a vektory matice A v závislosti na parametru p

(A) $p = 0{,}16$, $\lambda_1 = 1$, $\lambda_2 = 0{,}8$, $u_1 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$, $u_2 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$

(B) $p = 0{,}175$, $\lambda_1 = 0{,}95$, $\lambda_2 = 0{,}85$, $u_1 = \begin{pmatrix} 10 \\ 7 \end{pmatrix}$, $u_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

(C) $p = 0{,}135$, $\lambda_1 = 1{,}05$, $\lambda_2 = 0{,}75$, $u_1 = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$, $u_2 = \begin{pmatrix} 10 \\ 3 \end{pmatrix}$

Počáteční stav obou druhů můžeme napsat jako lineární kombinaci vlastních vektorů

$$X(0) = \begin{pmatrix} D_0 \\ K_0 \end{pmatrix} = a u_1 + b u_2$$

V čase n je stav obou populací

$$X(n) = A^n X(0) = A^n \begin{pmatrix} D_0 \\ K_0 \end{pmatrix} = A^n (a u_1 + b u_2) =$$

$$= a A^n u_1 + b A^n u_2 = a \lambda_1^n u_1 + b \lambda_2^n u_2\,,$$

Pro jednotlivé parametry to znamená

(A) $p = 0{,}16$, $\lambda_1^n = 1^n = 1$, $\lambda_2^n = 0{,}8^n \longrightarrow 0$, tedy

$$X_n \longrightarrow a u_1 = a \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$ Populace

se stabilizují.

(B) $p = 0{,}175$ $\lambda_1^n = 0{,}85^n \to 0$, $\lambda_2^n = 0{,}85^n \to 0$

proto $X(n) \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Populace mizí.

(C) $p = 0{,}135$ $\lambda_1^n = 1{,}05^n \to \infty$, $\lambda_2^n = 0{,}75^n \to 0$

tedy pro $a \neq 0$ populace roste

a poměr $D_n : K_n$

se blíží poměru dravců a kořisti ve vlastním vektoru u_1, tj. $10 : 9$.

(2) Leslieho populační model

Máme jeden druh s m věkovými skupinami. $x_i(n)$ je počet jedinců ve věkové skupině i v čase n. Model

$$X(n+1) = A\,X(n)$$

je dán maticí (tzv. Leslieho maticí)

$$A = \begin{pmatrix} f_1 & f_2 & f_3 & \cdots & f_{m-1} & f_m \\ \tau_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \tau_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \tau_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & \tau_{m-1} & 0 \end{pmatrix}$$

- číslo $f_i \geq 0$ určuje porodnost v i-té věkové skupině

$$x_1(n+1) = f_1 x_1(n) + f_2 x_2(n) + \dots + f_m x_m(n)$$

- τ_i je koeficient přežití z věku i do věku $i+1$. $\tau_i \in [0,1]$, $\tau_i = 1 -$ úmrtnost

$$x_{i+1}(n+1) = \tau_i x_i(n)$$

(3) Markovův proces

Systém se může nacházet v čase n v m různých stavech:

$p_i(n)$

je pravděpodobnost, že se nachází v čase n ve stavu i. Vektor

$$P(n) = \begin{pmatrix} p_1(n) \\ p_2(n) \\ \vdots \\ p_m(n) \end{pmatrix}$$

má vlastnost $\sum_{i=1}^{m} p_i(n) = 1$ a nazývá se pravděpodobnostní vektor.

V matici $A = (a_{ij})$ číslo $a_{ij} \in [0,1]$ udává pravděpodobnost, že se systém ze stavu j dostane do stavu i.

$$p_i(n+1) = a_{i1} p_1(n) + a_{i2} p_2(n) + \dots + a_{im} p_m(n)$$

To nám dává lineární proces

$$P(n+1) = A\,P(n),$$

který nazýváme Markovův proces.
Matice A má vlastnost, že součet prvků v každém sloupci je 1 a nazývá se stochastická nebo Markovova.

Perronova - Frobeniova teorie

Matice $A = (a_{ij})$ se nazývá pozitivní, jestliže $a_{ij} > 0$ pro všechna $1 \le i, j \le n$. Matice A se nazývá primitivní, jestliže nějaká její mocnina A^k je pozitivní.

Věta Perronova - Frobeniova

Nechť A je primitivní matice. Ze všech komplexních vlastních čísel má právě jedno největší absolutní hodnotu. Nazýváme jej dominantní. Toto vlastní číslo λ_1 má tyto vlastnosti:

(1) λ_1 je kladné reálné číslo

(2) geometrická násobnost vlastního čísla λ_1 je 1.

(3) K λ_1 existují vlastní vektory se všemi souřadnicemi kladnými

(4) Pokud je $\lambda_1 = 1$, systém určený modelem

$$X(n+1) = A\,X(n)$$

se stabilizuje a konverguje k násobku vlastního vektoru k λ_1.

(5) Pokud je $\lambda_1 < 1$, systém ve všech složkách konverguje exponenciálně k 0.

(6) Pokud je A Leslieho matice a $\lambda_1 > 1$, pak populace expanduje a poměry mezi jednotlivými populacemi se blíží k poměrům složek vlastního vektoru k vlastnímu číslu λ_1.

LEMMA: Markovova matice má dominantní vlastní číslo rovno rády jedné. Z každého počátečního stavu konverguje systém k vlastnímu pravděpodobnostnímu vektoru pro vlastní číslo 1.

S příkladem na Leslieho populační proces a na Markovovův proces se setkáte ve cvičení.