ÜBER DIE EIGENSCHAFT $(AB)^+ = B^+A^+$ DER MOORE-PENROSE-INVERSE

FRANK KLINKER

Satz 1.

1. Die *Moore-Penrose-Inverse* einer Matrix $A \in \mathbb{R}^{m\times n}$ ist die Matrix $A^+ \in \mathbb{R}^{n\times m}$, die durch die folgenden vier Eigenschaften eindeutig definiert ist.
   (i.1) $AA^+ = (AA^+)^T$
   (i.2) $A^+A = (A^+A)^T$
   (ii.1) $AA^+A = A$
   (ii.2) $A^+AA^+ = A^+$
2. Sei $A \in \mathbb{R}^{m\times n}$. Dann gibt es orthogonale Matrizen $U \in \mathbb{R}^{m\times m}$ und $V \in \mathbb{R}^{n\times n}$, sodass die Matrix $\Sigma := U^TAV$ eine reelle Diagonalmatrix mit Diagonalelementen $\Sigma_{ij} = \delta_{ij}\sigma_i$. Fordert man noch $\sigma_1 \geq \ldots \geq \sigma_r$ für $r = \min\{m, n\}$, so nennt man die Zerlegung
$$A = U\Sigma V^T$$
die *Singulärwertzerlegung* der Matrix A.
3. Ist $A = U\Sigma V^T$ die Singulärwertzerlegung der Matrix A mit der Diagonalmatrix $\Sigma_{ij} = \delta_{ij}\sigma_i$, so besitzt die Moore-Penrose-Inverse die Singulärwertzerlegung $A^+ = V\hat{\Sigma}U^T$ mit $\hat{\Sigma}_{ij} = \begin{cases} \delta_{ij}\sigma_i^{-1} & \text{für } \sigma_i \neq 0 \\ 0 & \text{sonst} \end{cases}$.

Aufgabe 2. Es seien $B \in \mathbb{R}^{m\times n}$ und $A \in \mathbb{R}^{n\times k}$ mit $rg(A) = rg(B) = n$ (insbesondere ist $n \leq \min\{m, k\}$. In diesem Fall gilt
$$(BA)^+ = A^+B^+\,.$$

Aufgabe 3. Es seien $B \in \mathbb{R}^{m\times n}$ und $A \in \mathbb{R}^{n\times k}$. Weiterhin sei mit $B^TB = \mathbb{1}$ oder $AA^T = \mathbb{1}$. In beiden Fällen gilt
$$(BA)^+ = A^+B^+\,.$$

Aufgabe 4. Im Allgemeinen gilt $(BA)^+ \neq A^+B^+$.

Bevor die Lösungen der Aufgaben folgen, geben wir noch eine Liste von hilfreichen Aussagen.

Lemma 5. Es sei $A \in \mathbb{R}^{m\times n}$ und $A^+ \in \mathbb{R}^{n\times m}$ sei sein Pseudoinverses. Dann gilt

Email: mail@frank-klinker.de.

(1) Es ist $(A^+)^+ = A$. Wenn A invertierbar ist, so gilt $A^+ = A^{-1}$.

(2) $rg(A) = rg(A^+A) = rg(A^+A) = rg(A^+) = sp(A^+A) = sp(A^+A)$, wobei sp die Spur bezeichnet.

(3) $(A^T)^+ = (A^+)^T$

(4) Ist A symmetrisch, also $A = A^T$, so ist A^+ ebenfalls symmetrisch und es gilt $A^+A = AA^+$

(5) $(A^TA)^+ = A^+(A^T)^+$

(6) $A^+ = (A^TA)^+A^T$ und $A^+ = A^T(AA^T)^+$

(7a) Es ist ist $m \geq n$ und $rg(A) = n$ genau dann wenn $A^+ = (A^TA)^{-1}A^T$ und $A^+A = E_n$.

(7b) Es ist $m \leq n$ und $rg(A) = m$ genau dann wenn $A^+ = A^T(AA^T)^{-1}$ und $AA^+ = E_m$.

Beweis von Lemma 5

Der erste Teil der Aussage (1) folgt aus der Tatsache, dass die Eigenschaften (i) und (ii) symmetrisch in A^+ und A sind. Ist A quadratisch und invertierbar so erfüllt A^{-1} die vier definierenden Eigenschaften (i) und (ii) was den zweiten Teil liefert.

Zum Beweis der Aussage (2) nutzt man die Singulärwertzerlegungen von A und A^+. und wir nehmen an, dass $rg(A) = r$ ist. Da U und V invertierbar sind gilt ebenfalls $rg(\Sigma) = r$ und per Definition $rg(\hat{\Sigma}) = rg(\Sigma)$. Dann also auch $rg(A^+) = rg(\hat{\Sigma}) = r$. Weiter ist $\Sigma\hat{\Sigma} = diag(\underbrace{1, \ldots, 1}_{r \text{ mal}}, 0 \ldots, 0) \in \mathbb{R}^{m\times m}$ und $\hat{\Sigma}\Sigma = diag(\underbrace{1, \ldots, 1}_{r \text{ mal}}, 0 \ldots, 0) \in \mathbb{R}^{n\times n}$, sowie $A^+A = V\hat{\Sigma}\Sigma V^T$ und $AA^+ = U\Sigma\hat{\Sigma}U^T$. Damit ist also ebenfalls $rg(A^+A) = rg(\hat{\Sigma}\Sigma) = r = rg(\Sigma\hat{\Sigma}) = rg(AA^+)$. Schließlich ist noch $sp(A^+A) = sp(AA^+) = sp(\Sigma\hat{\Sigma}) = r$.

Zum Beweis der Aussage (3) zeigt man, dass die Matrix $(A^+)^T$ die Eigenschaften (i) und (ii) der Pseudoinversen der Matrix A^T erfüllt. Damit ist dann nämlich $(A^T)^+ = (A^+)^T$.

(i.1): Zu zeigen ist $A^T(A^+)^T = (A^T(A^+)^T)^T$. Es ist

$$A^T(A^+)^T = (A^+A)^T \overset{\text{(i) für } A}{=} ((A^+A)^T)^T = (A^T(A^+)^T)^T\,.$$

(ii.1) Zu zeigen ist $A^T(A^+)^TA^T = A^T$. Es ist

$$A^T(A^+)^TA^T = (AA^+A)^T \overset{\text{(ii) für } A}{=} A^T\,.$$

Die Aussagen (i.2) und (ii.2) beweist man analog.

Aussage (4) sieht man wie folgt: Ist $A = A^T$ so gilt $(A^+)^T \overset{2.}{=} (A^T)^+ = A^+$, also ist auch A^+ symmetrisch. Nun ist

$$A^+A \overset{\text{(i) für } A}{=} (A^+A)^T = A^T(A^+)^T = AA^+ .$$

Zum Beweis der Aussage (5) zeigt man, dass die Matrix $A^+(A^T)^+$ die Eigenschaften (i) und (ii) der Pseudoinversen von A^TA erfüllt.

(i.1): Zu zeigen ist $(A^TA)(A^+(A^T)^+) = ((A^TA)(A^+(A^T)^+))^T$. Einerseits ist

$$\begin{aligned} A^T(AA^+)(A^T)^+ &\overset{\text{(i) für } A}{=} A^T(AA^+)^T(A^T)^+ \overset{3.}{=} (AA^+A)^T(A^+)^T \overset{\text{(ii) für } A}{=} A^T(A^+)^T \\ &= (A^+A)^T \end{aligned}$$

und andererseits

$$\begin{aligned} (A^TAA^+(A^T)^+)^T &\overset{3.}{=} A^+(AA^+)^TA \overset{\text{(i) für } A}{=} A^+AA^+A \overset{\text{(ii) für } A}{=} A^+A \\ &\overset{\text{(i) für } A}{=} (A^+A)^T \end{aligned}$$

(ii.1): Zu zeigen $(A^TA)(A^+(A^T)^+)(A^TA) = A^TA$. Es ist

$$\begin{aligned} A^TAA^+(A^T)^+A^TA &\overset{\text{(i) für } A}{=} A^T(AA^+)^T(A^T)^+A^TA \overset{3.}{=} (AA^+A)^T(A^+A)^TA \\ &\overset{\text{(ii) für } A}{=} (A^+)^T(A^+A)^TA = (A^+AA^+)^TA \\ &\overset{\text{(ii) für } A}{=} A^TA . \end{aligned}$$

Auch hier zeigt man (i.2) und (ii.2) analog.

Zum Beweis des ersten Teils der Aussage (6) zeigt man, dass $(A^TA)^+A^T$ die Eigenschaften (i) und (ii) der Pseudoinversen von A erfüllt.

(i.1): Zu zeigen ist $A((A^TA)^+A^T) = (A((A^TA)^+A^T))^T$. Die Matrix A^TA ist natürlich symmetrisch und deshalb gilt

$$(A(A^TA)^+A^T)^T = A((A^TA)^+)^TA^T \overset{4.}{=} A(A^TA)^+A^T$$

(ii.1): Zu zeigen ist $A((A^TA)^+A^T)A = A$. Es ist

$$\begin{aligned} A(A^TA)^+A^TA &\overset{5.}{=} AA^+(A^T)^+A^TA \overset{\text{(i) für } A}{=} (AA^+)^T(A^T)^+A^TA \\ &\overset{3.}{=} (AA^+)^T(A^+)^TA^TA = A^+A(A^+)^TA^TA \\ &\overset{\text{(ii) für } A}{=} (A^+)^TA^TA = (AA^+)^TA \overset{\text{(i) für } A}{=} AA^+A \\ &\overset{\text{(ii) für } A}{=} A \end{aligned}$$

Die Begründung für (i.2) und (ii.2) verläuft analog.

Zur Rückrichtung der Aussage (7a) bemerken wir zunächst, dass wegen $A^+A = \mathbb{1}$ die Matrix A ein Linksinverses besitzt und somit den vollen Rang $rg(A) = n$ hat. Ferner wird hier die weitere Voraussetzung nicht benötigt. Für die Hinrichtung sei nun umgekehrt der Rang von A maximal, also $rg(A) = n$. Dann hat A^TA ebenfalls den Rang n und ist somit invertierbar und wegen 1. gilt $(A^TA)^{-1} = (A^TA)^+$. Mit (6) folgt nun weiter $A^+ = (A^TA)^{-1}A^T$ und damit $A^+A = (A^TA)^{-1}A^TA = \mathbb{1}$.

Die Aussage (7b) beweist man analog.

Lösung der Aufgabe 2

Wir zeigen, dass die Matrix A^+B^+ die Eigenschaften (i) und (ii) der Pseudoinversen der Matrix BA erfüllt. Wegen der Bedingung an die Ränge der Matrizen A und B gilt wegen der Aussage 7. $B^+B = \mathbb{1}$ und $AA^+ = \mathbb{1}$.

(i.1): Zu zeigen ist $(BA)(A^+B^+) = ((BA)(A^+B^+))^T$. Es ist aber

$$(B\underbrace{AA^+}_{=\mathbb{1}}B^+)^T = (BB^+)^T \overset{\text{(i) für } B}{=} BB^+ = B\underbrace{AA^+}_{=\mathbb{1}}B^+ .$$

(ii.1): Zu zeigen ist $(BA)(A^+B^+)(BA) = BA$. Es ist jedoch

$$B\underbrace{AA^+}_{=\mathbb{1}}\underbrace{B^+B}_{=\mathbb{1}}A = BA .$$

Und auch hier beweist man (i.2) und (ii.2) analog.

Lösung der Aufgabe 3

Wie vorhin zeigen wir, dass die Matrix A^+B^+ die Eigenschaften (i) und (ii) der Pseudoinversen der Matrix BA erfüllt. Wegen der Voraussetzung $B^TB = \mathbb{1}$ ist

$$B^+ \overset{(7)}{=} (B^TB)^+B^T = B^T .$$

(i.1): Zu zeigen ist $(BA)(A^+B^+) = (BAA^+B^+)^T$. Es gilt

$$(BAA^+B^+)^T = (B^+)^T(AA^+)^TB^T \overset{B^+=B^T}{=} B(AA^+)^TB^+ \overset{\text{(i) für } A}{=} BAA^+B^+ .$$

(ii.2): Zu zeigen ist $(BA)(A^+B^+)(BA) = BA$. Es ist

$$BAA^+\underbrace{B^+B}_{=\mathbb{1}}A = BAA^+A \overset{\text{(ii) für } A}{=} BA .$$

Lösung der Aufgabe 4

Wähle $A = B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ mit $A^2 = A$. In diesem Fall gilt $(A^2)^+ \neq (A^+)^2$.

Es ist $A\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \sqrt{2}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ und $A\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 0$. Somit hat A mit $V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $U = \mathbb{1}$ und $\Sigma = \sqrt{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ die Singulärwertzerlegung

$$AV = U\Sigma$$

Damit und mit $\hat{\Sigma} = \frac{1}{2}\Sigma$ gilt

$$A^+ = V^T\hat{\Sigma}U = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix} = \frac{1}{2}A^T$$

Insgesamt ist somit

$$(A^2)^+ = A^+ = \frac{1}{2}A^T \neq \frac{1}{4}(A^T)^2 = (A^+)^2 \,.$$

Bemerkung 6. Die Aufgaben 2-4 zeigen, dass eine Charakterisierung der Matrizen A, B mit der Eigenschaft $(AB)^+ = B^+A^+$ ohne weitere Voraussetzung an die Matrizen nicht möglich zu sein scheint.