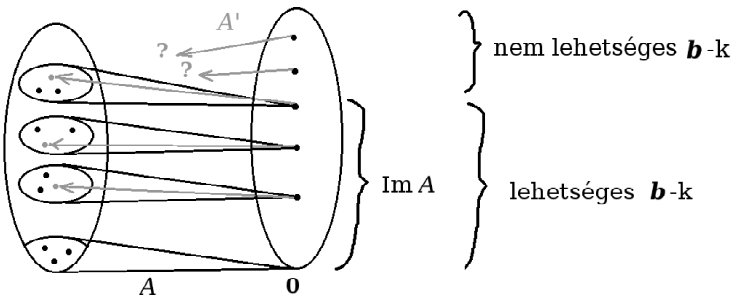
Moore–Penrose-féle pszeudoinverz

Ha nincs megoldása az $A\mathbf{x} = \mathbf{b}$ egyenletrendszernek, akkor a legjobban közelítő megoldást keressük: $\mathbf{x}_0$, amire $|A\mathbf{x}_0 - \mathbf{b}|$ minimális, azaz $A\mathbf{x}_0$ legyen $\operatorname{Im} A$-ban a $\mathbf{b}$-hez legközelebbi vektor ($\mathbf{b}$ merőleges vetülete $\operatorname{Im} A$-ra).



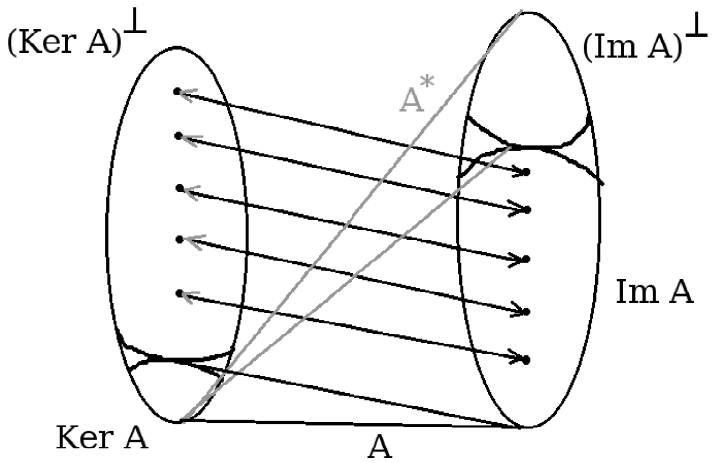
Olyan A' kell, amelyre AA' az $\operatorname{Im} A$-ra való merőleges vetítés.

Állítás. A egy merőleges vetítés mátrixa $\Leftrightarrow A^2 = A$, és A önadjungált.

Bizonyítás. $\Rightarrow$: Volt. $\Leftarrow$: $m_A(x) | x^2 - x = x(x-1) \Rightarrow$ minden sajátérték 0 és 1. Mivel A önadjungált is, unitérrel diagonalizálható, és a diagonális alakja $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$, tehát A valóban merőleges vetítés. $\square$

Def. $A \in \mathbb{C}^{m \times n}$-re $A^+ \in \mathbb{C}^{n \times m}$ az A **pszeudoinverze**, ha
(1) $AA^+A = A$;
(2) $A^+AA^+ = A^+$;
(3) $(AA^+)^* = AA^+$;
(4) $(A^+A)^* = A^+A$.

Megjegyzés. A fenti négy feltétel mindegyike teljesül, ha az alábbi módon választunk általánosított inverzet:



Állítás. Ⓑ $A\mathbf{x} = \mathbf{b}$-nek $A^+\mathbf{b}$ az egyik legjobban közelítő megoldása, azaz $|AA^+\mathbf{b} - \mathbf{b}| = \min\{\, |A\mathbf{x} - \mathbf{b}| \mid \mathbf{x} \in \mathbb{C}^n \,\}$.

Megjegyzés. Ebből következik, hogy ha van megoldás, akkor $A^+\mathbf{b}$ megoldás.

Bizonyítás. $\left.\begin{array}{l}(1) \Rightarrow\ AA^+AA^+ = AA^+ \\ (3) \Rightarrow\ AA^+ \text{ önadjungált}\end{array}\right\} \Rightarrow AA^+$ merőleges vetítés $\operatorname{Im} AA^+$-ra.
Mivel $\operatorname{Im} A \geq \operatorname{Im} AA^+ \geq \operatorname{Im} AA^+A = \operatorname{Im} A$, ez egyúttal az $\operatorname{Im} A = \{\, A\mathbf{x} \,|\, \mathbf{x} \in \mathbb{C}^m \,\}$-re való merőleges vetítés, így valóban $AA^+\mathbf{b}$ van legközelebb a $\mathbf{b}$-hez az $\operatorname{Im} A$ vektorai közül. □

Állítás. A pszeudoinverz egyértelmű.

Bizonyítás. Ha X és Y pszeudoinverz, akkor
$X \overset{(2)}{=} XAX \overset{(4)}{=} A^*X^*X \overset{(1)}{=} (A^*Y^*A^*)X^*X = (A^*Y^*)(A^*X^*)X \overset{(4)}{=} (YA)(XA)X \overset{(2)}{=} YAX$,
$Y \underset{(2)}{=} YAY \underset{(3)}{=} YY^*A^* \underset{(1)}{=} YY^*(A^*X^*A^*) = Y(Y^*A^*)(X^*A^*) \underset{(3)}{=} Y(AY)(AX) \underset{(2)}{=} YAX$.
Tehát $X = Y$. □

Pszeudoinverz kiszámítása

Lemma. Ⓑ $\operatorname{rang} A^*A = \operatorname{rang} A$ ($\Rightarrow \operatorname{rang} A = \operatorname{rang} A^* = \operatorname{rang} AA^*$)

Bizonyítás. A $\operatorname{rang} A^*A = \operatorname{rang} A$ egyenlőséghez elég: $\operatorname{Ker} A^*A = \operatorname{Ker} A$ (a dimenziótétel miatt, és mert mindkettő ugyanonnan képez).
$\geq$: nyilvánvaló.
$\leq$: $\mathbf{v} \in \operatorname{Ker} A^*A$
$\Rightarrow A^*A\mathbf{v} = \mathbf{0}$
$\Rightarrow \mathbf{0} = \mathbf{v}^*A^*A\mathbf{v} = (A\mathbf{v})^*(A\mathbf{v})$ □
$\Rightarrow A\mathbf{v} = 0$
$\Rightarrow \mathbf{v} \in \operatorname{Ker} A$.

A teljes rangú eset

Tétel. Legyen $A \in \mathbb{C}^{m\times n}$
1) Ha $\operatorname{rang} A = n \Rightarrow (A^*A)_{n\times n}$ invertálható, mert a lemma miatt n rangú, és $A^+ = (A^*A)^{-1}A^*$.
2) Ha $\operatorname{rang} A = m \Rightarrow (AA^*)_{m\times m}$ invertálható, mert a lemma miatt m rangú, és $A^+ = A^*(AA^*)^{-1}$.

Bizonyítás. Könnyen ellenőrizhetjük z pszeudoinverz definíciójában szereplő négy egyenlőséget. Pl. az első esetben $A' = (A^*A)^{-1}A^*$ bal inverze A-nak: $A'A = (A^*A)^{-1}A^*A = I$, és ebből (1), (2) és (4) közvetlenül kijön: $AA'A = AI = A$, $A'AA' = IA' = A'$, és $(A'A)^* = I^* = I = A'A$, míg
(3): $(AA')^* = (A(A^*A)^{-1}A^*)^* = A((A^*A)^{-1})^*A^* = A(A^*A)^{-1}A^* = AA'$; □

Általános eset

Tétel. Legyen $A \in \mathbb{C}^{m\times n}$, és $\operatorname{rang} A = r$.
Ekkor $\exists B \in \mathbb{C}^{m\times r}$ és $C \in \mathbb{C}^{r\times n}$ teljes rangú mátrixok, hogy $A = BC$.
És tetszőleges $A = BC$ felbontásra, ahol B és C teljes rangúak, $A^+ = C^+B^+$.

Bizonyítás. Az A mátrix oszlopaiból kiválasztható az oszloptér egy bázisa: ezek B oszlopai (r darab). C oszlopai azt mondják meg, hogy a B oszlopainak milyen lineáris kombinációja adja meg az A megfelelő oszlopait (ez $r \times m$-es). Így $BC = A$.

B és C egyszerre meghatározható egy Gauss-elimináció segítségével:

Hozzuk A-t elemi sorműveletekkel redukált lépcsős alakra!

B: A oszlopai közül az L-ben levő vezéregyesek oszlopainak megfelelő sorszámú oszlopok.
C: az L a $\mathbf{0}$ sorok nélkül.

Ezután $A^+ = C^+B^+$-re könnyen ellenőrizhetők a pszeudoinverzet definiáló (1), (2), (3), (4) tulajdonságok, felhasználva, hogy B^+ és C^+ bal-, illetve jobb inverz, amellett, hogy pszeudoinverz. □

Pl.: 1. Bontsuk fel az $A = \begin{bmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & -1 & 1 \\ 1 & 2 & 0 & -1 \end{bmatrix}$ mátrixot a fentiek szerint BC alakra, teljes rangú mátrixok szorzatára!

$$\begin{bmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & -1 & 1 \\ 1 & 2 & 0 & -1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 1 & -3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ebből leolvasható, hogy B az A első és harmadik oszlopából áll, és C a redukált lépcsős alak első két sorából:

$$B = \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{és} \quad \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -3 \end{bmatrix}.$$

2. Számítsuk ki az $A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -4 & 2 \end{bmatrix}$ mátrix pszeudoinverzét! Keressük meg az $A\mathbf{x} = \begin{bmatrix} 1.2 \\ -1.8 \end{bmatrix}$ megoldhatatlan egyenletrendszer legjobban közelítő megoldását.

$\operatorname{rang} A = 1$, tehát A nem teljes rangú.

$$\begin{bmatrix} 1 & 2 & -1 \\ -2 & -4 & 2 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow B = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad C = [\,1 \quad 2 \quad -1\,]$$

$$B^*B = [\,1 \quad -2\,] \begin{bmatrix} 1 \\ -2 \end{bmatrix} = [5], \quad B^+ = (B^*B)^{-1}B^* = \frac{1}{5}[\,1 \quad -2\,]$$

$$CC^* = [\,1 \quad 2 \quad -1\,] \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = [6], \quad C^+ = C^*(CC^*)^{-1} = \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$A^+ = C^+B^+ = \frac{1}{30} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} [\,1 \quad -2\,] = \frac{1}{30} \begin{bmatrix} 1 & -2 \\ 2 & -4 \\ -1 & 2 \end{bmatrix}$$

Ebből kiszámítható az egyenletrendszer egyik legjobban közelítő megoldása, $\mathbf{x}_0 = A^+ \begin{bmatrix} 1.2 \\ -1.8 \end{bmatrix} = \begin{bmatrix} 0.16 \\ 0.32 \\ -0.16 \end{bmatrix}$, és erre az $\mathbf{x}_0$-ra $A\mathbf{x}_0 = \begin{bmatrix} 0.96 \\ -1.92 \end{bmatrix}$.

SVD: szinguláris értékek szerinti felbontás (**S**ingular **V**alue **D**ecomposition)

Tudjuk: Ha A önadjungált, pozitív szemidefinit (a megengedő értelemben, tehát $\mathbf{x}^*A\mathbf{x} \geq 0\ \forall \mathbf{x}$), akkor $\exists U$ unitér, amelyre $A = UDU^*$, ahol $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, diagonális, és $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ (ugyanis A minden sajátértéke nemnegatív valós szám, és vehetjük az U oszlopaiban álló sajátvektorokat a sajátértékek csökkenő sorrendjében).

Ha A önadjungált, de nem pozitív szemidefinit, akkor legyen a D diagonális alak olyan, hogy $|\lambda_1| \geq \cdots \geq |\lambda_n|$. Az U oszlopaiban a negatív sajátértékű sajátvektorok előjelét ellenkezőjére változtatva kapjuk a szintén unitér U' mátrixot, amellyel

$$A = U' \begin{bmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{bmatrix} U^*.$$

Állítás. Ⓑ Legyen $A \in \mathbb{C}^{m\times n}$, $\operatorname{rang} A = r$.
Ekkor A^*A önadjungált, pozitív szemidefinit, és $\operatorname{rang} A^*A = r$.

Bizonyítás. $(A^*A)^* = A^*A$ miatt A önadjungált, és korábban beláttuk, hogy $\operatorname{rang} A^*A = \operatorname{rang} A$. Továbbá tetszőleges $\mathbf{x} \in \mathbb{C}^n$ vektorra $\mathbf{x}^*A^*A\mathbf{x} = (A\mathbf{x})^*(A\mathbf{x}) = |A\mathbf{x}|^2 \geq 0$, tehát A^*A pozitív szemidefinit. □

Def. $A \in \mathbb{C}^{m\times n}$ szinguláris értékei $\sigma_1 \geq \cdots \geq \sigma_r > 0$, ahol $\sigma_1^2, \ldots, \sigma_r^2$ az A^*A pozitív sajátértékei multiplicitással (ezekből éppen $r = \operatorname{rang} A$ darab van, mert $r = \operatorname{rang} A^*A$ miatt A^*A 0-hoz tartozó sajátaltere $n - r$ dimenziós).