Prof. Dr. Peter Kunkel
Fakultät für Mathematik und Informatik
Universität Leipzig

SS 2022

Ergänzung zur Vorlesung

Verallgemeinerte Inverse

1.6 Verallgemeinerungen und Nebenbedingungen

Zwei mögliche Verallgemeinerungen der Moore-Penrose-Pseudoinversen sind sofort denkbar. Zum einen kann man gewisse Axiome (1.13) weglassen. Zum anderen kann man auch das Problem (1.1) oder (1.4) in eine größere Klasse einbetten.

Satz 1.39
Sei $A \in \mathbb{C}^{m,n}$ und $A^- \in \mathbb{C}^{n,m}$. Es ist $x = A^- b$ für jedes $b \in \operatorname{bild} A$ Lösung von $Ax = b$ genau dann, wenn $AA^- A = A$ gilt, d. h. wenn A^- dem Axiom (1.13a) genügt. In diesem Fall läßt sich jede Lösung von $Ax = b$ in der Form

$$x = A^- b + (I - A^- A)v, \quad v \in \mathbb{C}^n \tag{1.56}$$

schreiben.

Beweis:
Sei $x = A^- b$ für jedes $b \in \operatorname{bild} A$ Lösung von $Ax = b$, d. h. sei $AA^- b = b$ für jedes $b \in \operatorname{bild} A$. Da jede Spalte von A in $\operatorname{bild} A$ liegt, folgt sofort $AA^- A = A$.

Sei umgekehrt $AA^- A = A$. Zu $b \in \operatorname{bild} A$ gibt es ein $u \in \mathbb{C}^n$ mit $b = Au$. Dann gilt $AA^- Au = Au$ bzw. $AA^- b = b$, d. h. $Ax = b$ für $x = A^- b$.

Für x aus (1.56) gilt mit $b = Au$

$$Ax = AA^- Au + A(I - A^- A)v = Au = b.$$

Jedes solche x ist also Lösung. Sei nun x_2 eine zweite Lösung neben $x_1 = A^- b$. Dann gilt $A(x_2 - x_1) = 0$. Wählt man nun $v = x_2 - x_1$, so gilt

$$x_2 = x_1 + (x_2 - x_1) = A^- b + (I - A^- A)v\,.$$

□

Lemma 1.40
Sei $A \in \mathbb{C}^{m,n}$ und genüge $A^- \in \mathbb{C}^{n,m}$ dem Axiom (1.13a), so gilt

$$\begin{array}{ll} \text{(a)} & \operatorname{rang} A^- \geq \operatorname{rang} A, \\ \text{(b)} & \operatorname{rang}(AA^-) = \operatorname{rang}(A^- A) = \operatorname{rang} A, \\ \text{(c)} & A^- A = I_n \text{ falls } \operatorname{rang} A = n, \\ \text{(d)} & AA^- = I_m \text{ falls } \operatorname{rang} A = m. \end{array} \tag{1.57}$$

Beweis:
Für (a) schließt man

$$\operatorname{rang} A = \operatorname{rang}(AA^-A) \le \operatorname{rang} A^-.$$

Für (b) schließt man

$$\begin{aligned}\operatorname{rang}(AA^-) \le \operatorname{rang} A = \operatorname{rang}(AA^-A) \le \operatorname{rang}(AA^-),\\ \operatorname{rang}(A^-A) \le \operatorname{rang} A = \operatorname{rang}(AA^-A) \le \operatorname{rang}(A^-A).\end{aligned}$$

Für (c) schließt man aus $A^-A \in \mathbb{C}^{n,n}$, $\operatorname{rang}(A^-A) = n$ und $(A^-A)^2 = A^-AA^-A = A^-A$, daß $A^-A = I_n$.

Für (d) schließt man aus $AA^- \in \mathbb{C}^{m,m}$, $\operatorname{rang}(AA^-) = m$ und $(AA^-)^2 = AA^-AA^- = AA^-$, daß $AA^- = I_m$.

□

Satz 1.41
Sei $A \in \mathbb{C}^{m,n}$ und genüge $A^- \in \mathbb{C}^{n,m}$ den Axiomen (1.13a,b), so gilt

$$\text{(1.58)}\qquad \begin{aligned}&\text{(a)}\qquad \operatorname{kern} A \oplus \operatorname{bild} A^- = \mathbb{C}^n,\\ &\text{(b)}\qquad \operatorname{bild} A \oplus \operatorname{kern} A^- = \mathbb{C}^m.\end{aligned}$$

Seien X_1, X_2 zwei solche Inversen von A mit

$$\text{(1.59)}\qquad \operatorname{kern} X_1 = \operatorname{kern} X_2, \quad \operatorname{bild} X_1 = \operatorname{bild} X_2,$$

so folgt $X_1 = X_2$, d. h. durch Vorgabe von Kern und Bild legt man eindeutig eine solche Inverse fest.

Beweis:
Ist $\operatorname{rang} A = r$, so liefert (1.13a,b) entsprechend (1.57)

$$\operatorname{rang} A = \operatorname{rang} A^- = \operatorname{rang}(A^-A) = \operatorname{rang}(AA^-) = r.$$

Insbesondere sind $P = A^-A$ und $Q = AA^-$ Projektoren auf $\operatorname{bild} A^-$ bzw. $\operatorname{bild} A$. Daraus folgt, daß $I - P = I - A^-A$ ein Projektor ist mit $\operatorname{bild}(I-P) \supseteq \operatorname{kern} A$. Da aber $\operatorname{rang}(I-P) = n - r = \dim \operatorname{kern} A$ ist, gilt $\operatorname{bild}(I-P) = \operatorname{kern} A$. Entsprechend folgt $\operatorname{bild}(I-Q) = \operatorname{kern} A^-$. Also

$$\begin{aligned}\mathbb{C}^n &= \operatorname{bild} P \oplus \operatorname{bild}(I-P) = \operatorname{bild} A^- \oplus \operatorname{kern} A,\\ \mathbb{C}^m &= \operatorname{bild} Q \oplus \operatorname{bild}(I-Q) = \operatorname{bild} A \oplus \operatorname{kern} A^-.\end{aligned}$$

Seien $X_1, X_2 \in \mathbb{C}^{n,m}$ jeweils mit (1.13a,b) sowie (1.59). Aus

$$\operatorname{kern} X_1 = \operatorname{kern} X_2 = \operatorname{bild}(I - AX_2)$$

schließt man $X_1(I - AX_2) = 0$ bzw. $X_1 = X_1AX_2$ oder $AX_1 = AX_1AX_2 = AX_2$. Aus

$$\operatorname{bild} X_1 = \operatorname{bild} X_2 = \operatorname{bild}(X_2A)$$

schließt man $X_1AX_2A = X_2A$ oder $X_1A = X_2A$. Damit gilt dann

$$X_1 = X_1AX_1 = X_2AX_1 = X_2AX_2 = X_2.$$

□

Satz 1.42
Sei $A \in \mathbb{C}^{m,n}$ und $A^- \in \mathbb{C}^{n,m}$. Es ist $x = A^-b$ für jedes $b \in \mathbb{C}^m$ Lösung von $\|Ax - b\|_2 = \min$ genau dann, wenn A^- den Axiomen (1.13a,c) genügt

Beweis:
Sei $x = A^-b$ für jedes $b \in \mathbb{C}^m$ Lösung von $\|Ax - b\|_2 = \min$, d. h. von $A^HAx = A^Hb$. Dann gilt für alle $b \in \mathbb{C}^m$

$$A^HAA^-b = A^Hb$$

und damit $A^HAA^- = A^H$. Daraus schließt man $(A^+)^HA^HAA^- = (A^+)^HA^H$ bzw. $AA^- = AA^+$, was sofort (1.13a,c) impliziert.

Sei nun $A^- \in \mathbb{C}^{n,m}$ mit (1.13a,c). Dann gilt für $b \in \mathbb{C}^m$ und $x = A^-b$

$$A^HAx = A^HAA^-b = A^H(A^-)^HA^Hb = (AA^-A)^Hb = A^Hb.$$

□

Satz 1.43
Sei $A \in \mathbb{C}^{m,n}$ und $A^- \in \mathbb{C}^{n,m}$. Es ist $x = A^-b$ für jedes $b \in \operatorname{bild} A$ Lösung von $Ax = b$ mit minimalem $\|x\|_2$ genau dann, wenn A^- den Axiomen (1.13a,d) genügt

Beweis:
Sei $x = A^-b$ für jedes $b \in \operatorname{bild} A$ die kürzeste Lösung von $Ax = b$. Dann gilt für alle $b \in \operatorname{bild} A$

$$A^-b = A^+b$$

bzw. für alle $u \in \mathbb{C}^n$

$$A^-Au = A^+Au,$$

d. h. $A^-A = A^+A$, was sofort (1.13a,d) impliziert.

Sei nun $A^- \in \mathbb{C}^{n,m}$ mit (1.13a,d) und sei $b = Au$. Dann ist $x = A^-b$ nach Satz 1.39 Lösung von $Ax = b$ und für jedes $v \in \operatorname{kern} A$ gilt

$$\begin{aligned}
\|x + v\|_2^2 &= x^Hx + x^Hv + v^Hx + v^Hv = \\
&= x^Hx + (A^-Au)^H + v^H(A^-Au) + v^Hv = \\
&= x^Hx + u^HA^- \underbrace{Av}_{=0} + \underbrace{v^HA^H}_{=0}(A^-)^Hu + v^Hv = \\
&= \|x\|_2^2 + \|v\|_2^2 \geq \|x\|_2^2.
\end{aligned}$$

□

Satz 1.44 (Bott/Duffin)
Sei $A \in \mathbb{C}^{n,n}$ und $L \in \mathbb{C}^{n,n}$ mit $M = ALL^+ + (I - LL^+)$ nichtsingulär. Dann hat das beschränkte Problem

$$Ax + y = b, \quad x \in \operatorname{bild} L, \quad y \in \operatorname{cobild} L \tag{1.60}$$

für jedes $b \in \mathbb{C}^n$ die eindeutige Lösung

$$x = A^L b, \quad y = (I - AA^L)b, \tag{1.61}$$

wobei die sog. Bott-Duffin-Inverse A^L von A gegeben ist durch

$$A^L = LL^+M^{-1}. \tag{1.62}$$

Diese erfüllt

$$\begin{array}{ll} \text{(a)} & LL^+ = A^L ALL^+ = LL^+AA^L, \\ \text{(b)} & A^L = LL^+A^L = A^L LL^+ \end{array} \tag{1.63}$$

und ist die eindeutige verallgemeinerte Inverse von $\tilde{A} = LL^+ALL^+$ mit $\operatorname{bild} A^L = \operatorname{bild} L$ und $\operatorname{kern} A^L = \operatorname{cobild} L$, die den Axiomen (1.13a,b) genügt.

Beweis:
Setzt man $z = x + y$, so ist $x = LL^+z$ und $y = (I - LL^+)z$ und (1.60) wird zu $Mz = b$ mit der einzigen Lösung $z = M^{-1}b$. Berücksichtigt man

$$(I - LL^+)M^{-1} = (M - ALL^+)M^{-1} = I - AA^L,$$

so ergibt sich daraus sofort (1.61). Multiplikation der obigen Beziehung von links mit LL^+ liefert $LL^+ = LL^+AA^L$. Aus $MLL^+ \overset{\text{Def. } M}{=} ALL^+$ und $A^L M \overset{(1.62)}{=} LL^+$ folgt

$$LL^+ = LL^+LL^+ = A^L M M^{-1} ALL^+ = A^L ALL^+.$$

An (1.62) erkennt man sofort, daß $LL^+A^L = A^L$ ist. Schließlich gilt

$$A^L(I - LL^+) = A^L(M - ALL^+) = A^L M - A^L ALL^+ = LL^+ - LL^+ = 0$$

und damit $A^L = A^L LL^+$. Aus (1.63) erhält man nun

$$\tilde{A}A^L\tilde{A} = \overbrace{LL^+A\underbrace{LL^+A^L}_{=A^L}}^{=LL^+} LL^+ALL^+ = \tilde{A}, \;\; A^L\tilde{A}A^L = \overbrace{\underbrace{A^L LL^+}_{=A^L} ALL^+}^{=LL^+} A^L = A^L.$$

Daß $\operatorname{bild} A^L = \operatorname{bild} L$, entnimmt man sofort (1.62). Außerdem gilt

$$\begin{aligned} \operatorname{kern} A^L &= \operatorname{kern}(A^L LL^+) = \operatorname{cobild}(LL^+(A^L)^H) \supseteq \\ &\supseteq \operatorname{cobild}(LL^+) = \operatorname{cobild}(LL^+M^{-1}) = \operatorname{cobild} A^L. \end{aligned}$$

Da bei einer quadratischen Matrix Dimension von Kern und Kobild gleich sind, gilt sogar

$$\operatorname{kern} A^L = \operatorname{cobild} A^L = \operatorname{cobild} L.$$

∎

Satz 1.45
Sei $A \in \mathbb{C}^{m,n}$, $b \in \mathbb{C}^m$, $C \in \mathbb{C}^{p,n}$, $d \in \mathbb{C}^p$ und $T^H T = I$ mit $\operatorname{bild} T = \operatorname{kern} C$. Das Minimierungsproblem

$$\|x\|_2 = \min \ \text{ s. t. } \|Ax - b\|_2 = \min \ \text{ s. t. } \|Cx - d\|_2 = \min \tag{1.64}$$

besitzt die eindeutige Lösung

$$x = C^+ d + T(AT)^+(b - AC^+ d). \tag{1.65}$$

Beweis:
Alle Lösungen von $\|Cx - d\|_2 = \min$ haben die Form

$$x = C^+ d + Tw.$$

Der Vektor w ist dann aus dem Problem

$$\|C^+ d + Tw\|_2 = \min \ \text{ s. t. } \|AC^+ d + ATw - b\|_2 = \min$$

zu bestimmen. Wegen

$$\begin{aligned} \|C^+ d + Tw\|_2^2 &= d^H (C^+)^H C^+ d + d^H \underbrace{(C^+)^H T}_{=0} w + w^H \underbrace{T^H C^+}_{=0} d + w^H T^H T w = \\ &= d^H (C^+)^H C^+ d + w^H w = \|C^+ d\|_2^2 + \|w\|_2^2 \end{aligned}$$

ist dieses äquivalent zu

$$\|w\|_2 = \min \ \text{ s. t. } \|ATw - (b - AC^+ d)\|_2 = \min$$

mit der Lösung

$$w = (AT)^+(b - AC^+ d),$$

was dann (1.65) liefert.

$\square$

Bemerkung 1.46
Setzt man mit den Bezeichnungen des vorigen Satzes $E = I - C^+ C = TT^H$, so gilt mit

$$(AT)^+ = (T^H A^H AT)^+ T^H A^H \tag{1.66}$$

nach (1.18h) und

$$T(T^H A^H AT)^+ T^H = (EA^H AE)^+, \tag{1.67}$$

daß x in (1.65) dargestellt werden kann als

$$x = C^+ d + (EA^H AE)^+ (A^H b - A^H AC^+ d). \tag{1.68}$$

Dies kann dazu benutzt werden zu zeigen, daß man x auch auffassen kann als erste Komponente von

$$(1.69) \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A^H A & C^H \\ C & 0 \end{bmatrix} \begin{bmatrix} A^H b \\ d \end{bmatrix}.$$

vgl. Normalgleichung bzw. Kuhn-Tucker-Matrix

Setzt man

$$(1.70) \qquad A^C = (AE)^+ = T(AT)^+ = (EA^H AE)^+ EA^H = (EA^H AE)^+ A^H,$$

letzteres wegen $T^H E = T^H$ und (1.67), so ergibt sich

$$(1.71) \quad x = (AE)^+ b + (I - (AE)^+ A)C^+ d = A^C b + (I - A^C A)C^+ d,$$

vgl. (1.56)

weswegen A^C auch als bzgl. C eingeschränkte Pseudoinverse von A bezeichnet wird.

$\square$