Universität Dortmund
Fachbereich Mathematik
Prof. Dr. H. M. Möller
Dipl. Math. Katrin Siemko

Dortmund, den 22. 1. 2008

Numerische Mathematik I

Ergänzung zum Abschnitt 2.6 (Singulärwertzerlegung)
– kein prüfungsrelevanter Stoff –

Die Existenz der Singulärwertzerlegung wurde in der Vorlesung konstruktiv bewiesen (Beweis von Satz 11). Hier wollen wir eine alternative Konstruktion zeigen. Man findet diese Konstruktion auch in einigen anderen Büchern, z.B. im Buch von G. Golub und C.F. van Loan, Matrix Computations, The John Hopkins University Press, 1996.

Bemerkung. Die zur Euklidischen Norm $\| \, . \, \|_2$ gehörende lub_2-Norm ist (auch im Fall rechteckiger Matrizen A) die Spektralnorm

$$M_2(A) := \max\{ \sqrt{\lambda} \mid \lambda \text{ Eigenwert von } A^H A \}.$$

Damit ist $M_2(A)$ der größte Singulärwert der Matrix A.

Als Alternative zu der in der Vorlesung angegebenen Methode zur Berechnung der Singulärwertzerlegung hat man folgendes Verfahren.

Ein Verfahren zur Singulärwertzerlegung.
Eingabe: Matrix $A \in \mathbb{C}^{m \times n}$.
Ausgabe: Unitäre Matrizen $U \in \mathbb{C}^{m \times m}$ und $V \in \mathbb{C}^{n \times n}$ und Matrix $\Sigma := (s_{ij}) \in \mathbb{C}^{m \times n}$ mit

$$A = U \Sigma V^H \qquad \text{(Singulärwertzerlegung von } A\text{)}.$$

Dabei sind die Einträge s_{ij} von Σ bestimmt durch $s_{ii} := \sigma_i, \; i = 1, \ldots, r$, und $s_{ij} = 0$ für alle anderen Einträge von Σ. Und $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$.

Rechnung: Setze $\sigma_1 := M_2(A)$. Wegen $M_2(A) = \max_{\|x\|_2 = 1} \|Ax\|_2$ existiert ein $x \in \mathbb{C}^n$ mit $\|x\|_2 = 1$ und $\|Ax\| = \sigma_1$. Der Vektor $y := \frac{1}{\sigma_1} Ax \in \mathbb{C}^m$ erfüllt dann auch $\|y\|_2 = 1$ und es gilt

$$Ax = \sigma_1 y.$$

Man ergänze $\{x\}$ zu einer ONB von $\mathbb{C}^n$ und $\{y\}$ zu einer des $\mathbb{C}^m$,

$$V := \Big(x, x_2, \ldots, x_n \Big) \in \mathbb{C}^{n \times n}, \quad U := \Big(y, y_2, \ldots, y_m \Big) \in \mathbb{C}^{m \times m}.$$

Damit ergibt sich

$$A_1 := U^H A V = \begin{pmatrix} \sigma_1 & w^H \\ 0 & B \end{pmatrix}$$

mit einem Vektor $w \in \mathbb{C}^{n-1}$ und $B \in \mathbb{C}^{(m-1)\times(n-1)}$. Der Vektor w ist der Nullvektor (wird unten gezeigt).

Man erhält daher

$$A_1 = U^H A V = \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix}$$

und kann dann auf die kleinere Matrix $B \in \mathbb{C}^{(m-1)\times(n-1)}$ die selbe Technik anwenden,

$$B_1 = U_1^H B V_1 = \begin{pmatrix} \sigma_2 & 0 \\ 0 & C \end{pmatrix}.$$

Vergrößert man die Matrizen $U_1 \in \mathbb{C}^{(m-1)\times(m-1)}$ und $V_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ durch Rändern,

$$\tilde{U_1} := \begin{pmatrix} 1 & 0 \\ 0 & U_1, \end{pmatrix} \in \mathbb{C}^{m\times m}, \quad \tilde{V_1} := \begin{pmatrix} 1 & 0 \\ 0 & V_1, \end{pmatrix} \in \mathbb{C}^{n\times n},$$

dann bekommt man

$$\tilde{U}_1^U U^H A V \tilde{V_1} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & C \end{pmatrix}.$$

Dann verfährt man genauso mit der Matrix C usw. bis schließlich die Matrix $\Sigma = diag(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$ gefunden ist und man durch Zusammenfassen des Produkts der verschiedenen unitären U^H zu einem unitären U^H und der verschiedenen V zu einem V die Zerlegung von A

$$U^H A V = \Sigma$$

bekommt, aus der dann die Singulärwertzerlegung $A = U \Sigma V^H$ resultiert. □

Wir wollen $w = 0$ zeigen. Da U und V unitär sind, folgt nach Satz 10

$$M_2(A_1) = M_2(U^H A V) = M_2(U^H) M_2(A) M_2(V) = M_2(A) = \sigma_1.$$

Direkte Rechnung ergibt für den Vektor $\begin{pmatrix} \sigma_1 \\ w \end{pmatrix}$

$$\begin{aligned}
\left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 &= (\sigma_1, w^H) \begin{pmatrix} \sigma_1 & 0 \\ w & B^H \end{pmatrix} \begin{pmatrix} \sigma_1 & w^H \\ 0 & B \end{pmatrix} \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \\
&= (\sigma_1^2 + w^H w, w^H B^H) \begin{pmatrix} \sigma_1^2 + w^H w \\ Bw \end{pmatrix} = \left(\sigma_1^2 + \|w\|_2^2\right)^2 + \|Bw\|_2^2 \\
&\geq \left(\sigma_1^2 + \|w\|_2^2\right) \cdot \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 + 0 \geq \sigma_1^2 \cdot \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2.
\end{aligned}$$

Ist hierbei $w \neq 0$, dann ist das letzte $\geq$-Symbol durch $>$ zu ersetzen. Andererseits gilt

$$\left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 \leq M_2(A_1)^2 \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 = \sigma_1^2 \cdot \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2.$$

Man erhält also eine Ungleichungskette

$$\sigma_1^2 \cdot \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 \;\geq\; \left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 \;\geq\; \sigma_1^2 \cdot \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2.$$

Der Vergleich ergibt, dass überall das Gleichheitszeichen stehen muss, insbesondere also $w = 0$.

Die Pseudo-Inverse A^+ einer reellen Matrix $A \in \mathbb{R}^{m \times n}$ ist laut Vorlesung dadurch charakterisiert, dass A^+b unter allen Lösungen x^* des Minimierungsproblems

$$\|Ax^* - b\|_2^2 = \min\{\ \|Ax - b\|_2^2 \mid x \in \mathbb{R}^n\ \}$$

die kleinste (Euklidische) Norm besitzt. Eine äquivalente Charakterisierung ist dadurch gegeben, dass das Paar A, A^+ die vier Moore-Penrose Gleichungen erfüllt.

Die Moore-Penrose Gleichungen lassen sich natürlich auch komplex formulieren,

$$AB = (AB)^H, \quad BA = (BA)^H, \quad ABA = A, \quad BAB = B.$$

Damit bekommt man auch für $A \in \mathbb{C}^{m \times n}$ eine Pseudo-Inverse A^+. In der Vorlesung wurde gezeigt, wie man aus der Singulärwertzerlegung einer Matrix $A \in \mathbb{R}^{m \times n}$ die Pseudo-Inverse A^+ erhalten kann. Dies erfordert die Berechnung sämtlicher Eigenwerte der Matrix $A^T A$. Im komplexen Fall muss man entsprechend die Eigenwerte von $A^H A$ nehmen und ansonsten analog verfahren.

Die Bestimmung der Eigenwerte einer Matrix ist in der Regel sehr aufwändig. Hier wollen wir daher zeigen, wie man mit Methoden der linearen Algebra die Pseudo-Inverse berechnen kann. Um die Allgemeinheit des Verfahrens zu betonen, werden wir mit K eine beliebigen Körper mit $\mathbb{Q} \subseteq K \subseteq \mathbb{C}$ bezeichnen und von einer Matrix $A \in K^{m \times n}$ ausgehen und eine Matrix $B \in K^{n \times m}$ konstruieren, die mit A zusammen ein Paar bildet, das die vier Moore-Penrose Gleichungen erfüllt.

Nebenbemerkung. Das innere Produkt in K^n ist definiert durch $\langle x, y\rangle := x^T y$ wenn $K \subseteq \mathbb{R}$ und durch $\langle x, y\rangle := x^H y$ sonst. Aus Bequemlichkeit schreiben wir in beiden Fällen $x^H y$ für $\langle x, y\rangle$. Analog für Matrizen A^H statt A^T, selbst wenn $A \in \mathbb{R}^{m\times n}$.

Zunächst der aus der linearen Algebra bekannte

Nullraum-Algorithmus.
Eingabe: Matrix $A \in K^{m\times n}$.
Ausgabe: Eine Basis des K-Vektorraums $\{x \in K^n \mid Ax = 0\}$.

Rechnung:
Bringe A mit elementaren Zeilenoperationen auf Zeilenstufenform

$$\tilde{A} = \begin{pmatrix} 0 & \cdots & 0 & a_{1j_1} & \cdots & & & \cdots & a_{1n} \\ 0 & \cdots & \cdots & 0 & a_{2j_2} & \cdots & & \cdots & a_{2n} \\ \vdots & & & & \ddots & \ddots & & & \vdots \\ 0 & \cdots & & & \cdots & 0 & a_{rj_r} & \cdots & a_{rn} \\ 0 & \cdots & & & & & & \cdots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \cdots & & & & & & \cdots & 0 \end{pmatrix}$$

mit den Pivots $a_{\nu j_\nu} \neq 0,\ \nu = 1, \ldots, r$.

Sei $\{k_1, \ldots, k_{n-r}\} \cup \{j_1, \ldots, j_r\} = \{1, 2, \ldots, n\}$. Dann wählt man für jedes $s \in \{1, \ldots, n-r\}$ ein $v^{(s)} \in K^n$ so, dass

$$v_s^{(s)} := 1, \quad v_{k_\nu}^{(s)} := 0 \ \text{ für } \ k_\nu \in \{k_1, \ldots, k_{n-r}\} \setminus \{s\}.$$

Damit ist die ν-te Gleichung von $\tilde{A}v^{(s)} = 0$

$$\sum_{i=1}^{r} a_{\nu j_i} v_{j_i}^{(s)} + a_{\nu s} = 0, \quad \nu = 1, \ldots, r. \tag{1}$$

Die Koeffizientenmatrix $(a_{\nu j_i})$ ist eine obere Dreicksmatrix mit den Pivots auf der Diagonalen. Daher ist das Gleichungssystem (1) eindeutig für $v_{j_1}^{(s)}, \ldots, v_{j_r}^{(s)}$ lösbar. Damit sind alle Komponenten von $v^{(s)}$ bestimmt und es gilt $\tilde{A}v^{(i)} = 0$.

Die Vektoren $v^{(1)}, \ldots, v^{(n-r)}$ sind lin. unabhängig, denn die Zeilen $k_1, ..., k_{n-r}$ der Matrix mit diesen $n-r$ Vektoren als Spalten bilden die $(n-r)\times(n-r)$-Einheitsmatrix. Nach Konstruktion liegen $v^{(1)}, \ldots, v^{(n-r)}$ im Nullraum von $\tilde{A}$, der ja auch Nullraum von A ist (Gaußelimination!). Der Nullraum von A hat die Dimension $n - r$. Also haben wir mit den Vektoren $v^{(1)}, \ldots, v^{(n-r)}$ eine Basis des Nullraums von A konstruiert.

Algorithmus zur Berechnung der Pseudo-Inversen
Eingabe: Matrix $A \in K^{m\times n}$.
Ausgabe: Pseudo-Inverse $A^+ \in K^{n\times m}$.

Schritt 1:
$\{c_1,\ldots,c_{n-r}\} \subset K^n$ sei eine Basis des Nullraums von A. Die Matrix mit den Zeilen $c_1^H,\ldots,c_{n-r}^H$ werde mit C bezeichnet. (Es gilt $C \in K^{(n-r)\times n}$.) Mit der $m\times m$-Einheitsmatrix E_m und einer Nullmatrix $0 \in K^{(n-r)\times m}$ bilde man

$$\mathcal{A} := \begin{pmatrix} A & E_m \\ C & 0 \end{pmatrix} \in K^{(m+n-r)\times(n+m)}.$$

Schritt 2:
Man wende auf die Matrix $\mathcal{A}$ vollständige Gaußelimination (mit Zeilenvertauschung) für die ersten n Spalten an, d.h., man breche die Gaußelimination ab, sobald man die ersten n Spalten von $\mathcal{A}$ zu den ersten n Einheitsspalten gemacht hat. Das gibt eine Matrix

$$\tilde{\mathcal{A}} := \begin{pmatrix} E_n & B \\ 0 & D \end{pmatrix}.$$

Hierbei ist $B \in K^{n\times m}$ und $D \in K^{(m-r)\times m}$.

Schritt 3:
Man berechne die Pseudoinverse D^+ von D mittels

$$D^+ = D^H\left(DD^H\right)^{-1}$$

und gebe

$$\tilde{B} := B - BD^+D$$

als Pseudoinverse von A aus. □

Zur Durchführbarkeit des Verfahrens ist zu zeigen, dass
(i) $\tilde{\mathcal{A}}$ existiert, d.h., man kann mit Gaußelimination tatsächlich auf die Gestalt von $\tilde{\mathcal{A}}$ kommen.
(ii) D hat vollen Zeilenrang ($\Rightarrow D^+ = D^H\left(DD^H\right)^{-1}$).
Zur Korrektheit ist zu zeigen, dass am Ende schließlich $A^+ = \tilde{B}$ gilt.

Zu (i):
Die Matrix $\begin{pmatrix} A \\ C \end{pmatrix}$ hat $m+n-r$ Zeilen. Die ersten m spannen einen r-dimensionalen Raum auf, die letzten $n-r$ einen dazu orthogonalen $n-r$-dimensionalen Raum. Alle Zeilen zusammen spannen also einen n-dimensionalen

Raum auf. Die Matrix $\begin{pmatrix} A \\ C \end{pmatrix} \in K^{(m+n-r)\times n}$ hat somit vollen Spaltenrang. Man kann sie daher mit Gaußelimimation auf eine Gestalt $\begin{pmatrix} E_n \\ 0 \end{pmatrix}$ bringen und folglich auch $\mathcal{A}$ auf $\tilde{\mathcal{A}}$.

Zu (ii):
Die Matrix $\mathcal{A}$ hat den vollen Zeilenrang $m+n-r$, denn die ersten m Zeilen sind linear unabhängig (man betrachte die letzten m Komponenten!) und die letzten $n-r$ sind auch untereinander linear unabhängig und stehen orthogonal auf dem von den ersten m Zeilen aufgespannten Raum. Daher hat auch $\tilde{\mathcal{A}}$ vollen Zeilenrang. Insbesondere sind die letzten $m-r$ Zeilen linear unabhängig. Also gilt $\operatorname{Rang}(D) = m-r$.

Zum Nachweis der Korrektheit formen wir zuerst etwas um. Die Matrix $\mathcal{A}$ aus Schritt 1 ergab durch elementare Zeilenumformungen (und eventuelle Zeilenvertauschungen) die Matrix $\tilde{\mathcal{A}}$. Diese Umformungen kann man auch darstellen durch eine Multiplikation mit einer invertierbaren Matrix M_1,

$$M_1\mathcal{A} = \tilde{\mathcal{A}}.$$

Mit

$$M_2 := \begin{pmatrix} E_n & -BD^+ \\ 0 & E_{m-r} \end{pmatrix}$$

ist eine obere Dreiecksmatrix gegeben, die invertierbar ist (Auf Hauptdiagonale nur Einsen). Dann gibt die (invertierbare) Matrix $M := M_2M_1$

$$M\begin{pmatrix} A & E_m \\ C & 0 \end{pmatrix} = M_2\begin{pmatrix} E_n & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} E_n & \tilde{B} \\ 0 & D \end{pmatrix}.$$

Dabei sind die Zeilen von C orthogonal zu denen von A, also $AC^H = 0$, und auch die Zeilen von D sind orthogonal zu denen von $\tilde{B}$, denn

$$\tilde{B}D^H = (B - BD^+D)D^H = BD^H - BD^H\left(DD^H\right)^{-1}DD^H = 0.$$

Satz. Sei $A \in K^{m\times n}$ eine Matrix mit $\operatorname{Rang}(A) = r$. Die Zeilen von $C \in K^{(n-r)\times n}$ seien orthogonal zu denen von A und die Matrix $\begin{pmatrix} A \\ C \end{pmatrix}$ habe Rang n. Ferner seien $\tilde{B} \in K^{n\times m}$ und $D \in K^{(m-r)\times m}$ so, dass die Zeilen von D orthogonal zu denen von $\tilde{B}$ sind. $M \in K^{(m+n-r)\times(m+n-r)}$ sei invertierbar und es gelte

$$M\begin{pmatrix} A & E_m \\ C & 0 \end{pmatrix} = \begin{pmatrix} E_n & \tilde{B} \\ 0 & D \end{pmatrix}. \tag{2}$$

Dann folgt $\tilde{B} = A^+$.

Beweis. C und D haben jeweils vollen Zeilenrang (s. Durchführbarkeit des Algorithmus'). Daher sind CC^H und DD^H invertierbar und man erhält $C^+ = C^H\left(CC^H\right)^{-1}$ sowie $D^+ = D^H\left(DD^H\right)^{-1}$.

Wir partitionieren M in $M = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}$ mit

$$U_1 \in K^{n\times m},\ U_2 \in K^{n\times(n-r)},\ U_3 \in K^{(m-r)\times m},\ U_4 \in K^{(m-r)\times(n-r)}.$$

Dann gilt

$$\begin{array}{lll}
U_1A + U_2C = E_n & \Rightarrow U_1AC^H + U_2CC^H = C^H & \Rightarrow U_2 = C^+, \\
U_1E_m + 0 = \tilde{B} & \Rightarrow U_1 = \tilde{B}, & \\
U_3A + U_4C = 0 & \Rightarrow U_3AC^H + U_4CC^H = 0 & \Rightarrow U_4 = 0, \\
U_3E_m + 0 = D & \Rightarrow U_3 = D. &
\end{array}$$

Also $M = \begin{pmatrix} \tilde{B} & C^+ \\ D & 0 \end{pmatrix}$. Es folgt dann aus (2)

$$\tilde{B}A + C^+C = E_n.$$

Weil C^+C und E_n Hermitesch sind, gilt das auch für $\tilde{B}A$. Die erste der vier Moore-Penrose Gleichungen ist damit bewiesen.

Da M invertierbar ist, bekommt man auf analoge Art aus

$$M^{-1}\begin{pmatrix} E_n & \tilde{B} \\ 0 & D \end{pmatrix} = \begin{pmatrix} A & E_m \\ C & 0 \end{pmatrix} \tag{3}$$

die Darstellung $M^{-1} = \begin{pmatrix} A & D^+ \\ C & 0 \end{pmatrix}$ und daher aus (3)

$$A\tilde{B} + D^+D = E_m.$$

Insbesondere folgt dann, dass $A\tilde{B}$ Hermitesch ist, d.h., die zweite Moore-Penrose Gleichung.

Wegen $\tilde{B}A + C^+C = E_n$ und $A\tilde{B} + D^+D = E_m$ folgt schließlich wg. $AC^H = 0$

$$A\tilde{B}A = A(E_n - C^+C) = A - AC^H\left(CC^H\right)^{-1}C = A$$

und wegen $\tilde{B}D^H = 0$

$$\tilde{B}A\tilde{B} = \tilde{B}(E_m - D^+D) = \tilde{B} - \tilde{B}D^H\left(DD^H\right)^{-1}D = \tilde{B}.$$

Dies sind die fehlenden zwei Moore-Penrose Gleichungen für das Matrizenpaar $A, \tilde{B}$. $\square$

Beispiel 1. Die Matrix $A := (-1, 1, 0) \in K^{1\times 3}$ hat vollen Zeilenrang. Also gilt $A^+ = A^H \cdot (2)^{-1} = (-\frac{1}{2}, \frac{1}{2}, 0)^H$. Wir wollen hier aber zeigen, dass man dieses Resultat auch mit dem neuen Algorithmus bekommt.

Im ersten Schritt wird eine Basis des Nullraums von A bestimmt. $(1, 1, 0)^H$ und $(0, 0, 1)^H$ bilden eine Basis. Daher ist hier – mit der "Einheitsmatrix" $E_1 = (1)$

$$\mathcal{A} = \left(\begin{array}{ccc|c} -1 & 1 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Die Gaußelimination in Schritt 2 führt zu

$$\tilde{\mathcal{A}} = \left(\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Hier tritt kein D auf, also keine Modifikation von B erforderlich, $\tilde{B} = B$. Man liest ab:

$$A^+ = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}.$$

Beispiel 2. Wir betrachten für $\varepsilon \neq 1$

$$A_\varepsilon := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & \varepsilon \end{pmatrix}.$$

Diese Rang-2-Matrix hat als Nullraum den von $(1, -1, 0)^H$ aufgespannten Raum. Daher gibt Schritt 1

$$\mathcal{A}_\varepsilon = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & \varepsilon & 0 & 0 & 1 \\ \hline 1 & -1 & 0 & 0 & 0 & 0 \end{array}\right).$$

In Schritt 2 bekommt man mit einiger (aber elementarer) Rechnung

$$\tilde{\mathcal{A}}_\varepsilon = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ 0 & 1 & 0 & \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ 0 & 0 & 1 & \frac{-1}{\varepsilon-1} & 0 & \frac{1}{\varepsilon-1} \\ \hline 0 & 0 & 0 & -1 & 1 & 0 \end{array}\right).$$

Mit der Pseudo-Inversen $D^+ = (-\frac{1}{2}, \frac{1}{2}, 0)^H$ von $D = (-1, 1, 0)$ bekommt man dann in Schritt 3

$$\begin{aligned} A_\varepsilon^+ &= \begin{pmatrix} \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ \frac{-1}{\varepsilon-1} & 0 & \frac{1}{\varepsilon-1} \end{pmatrix} - \begin{pmatrix} \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ \frac{\varepsilon}{2\varepsilon-2} & 0 & \frac{-1}{2\varepsilon-2} \\ \frac{-1}{\varepsilon-1} & 0 & \frac{1}{\varepsilon-1} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} (-1, 1, 0) \\ &= \begin{pmatrix} \frac{\varepsilon}{4\varepsilon-4} & \frac{\varepsilon}{4\varepsilon-4} & \frac{-1}{2\varepsilon-2} \\ \frac{\varepsilon}{4\varepsilon-4} & \frac{\varepsilon}{4\varepsilon-4} & \frac{-1}{2\varepsilon-2} \\ \frac{-1}{2\varepsilon-2} & \frac{-1}{2\varepsilon-2} & \frac{1}{\varepsilon-1} \end{pmatrix}. \end{aligned}$$

Beispiel 3. Sei $A \in K^{n\times n}$ invertierbar. Die Pseudo-Inverse ist dann natürlich A^{-1}. Man berechnet etwa A^{-1} dadurch, dass man auf (A, E_n) vollständige Gaußelimination anwendet und (E_n, A^{-1}) erhält. Das ist aber exakt dasselbe beim neuen Algorithmus:

Der Nullraum von A wird durch keinen Vektor $\neq 0$ aufgespannt. Die Matrix $\mathcal{A}$ in Schritt 1 ist daher

$$\mathcal{A} = \Big(A, E_n\Big).$$

Gaußelimination in Schritt 2 gibt dann

$$\tilde{\mathcal{A}} = \Big(E_n, A^{-1}\Big).$$

Schritt 3 ist wieder überflüssig, weil kein D da ist. A^{-1} ist die gesuchte Pseudo-Inverse.

Im Normalfall ist nicht bekannt, ob $A \in K^{m\times n}$ überhaupt einen Nullraum $\neq \{0\}$ besitzt, geschweige denn, wie eine Basis des Nullraums aussieht. Man kann aber die Schritte 1 und 2 des Algorithmus mischen. Also erst mit der Matrix $(A, E_m) \in K^{m\times(n+m)}$ starten und auf Zeilenstufenform $(\hat{A}, \hat{E}_m)$ für die ersten n Spalten bringen. Aus der Zeilenstufenform berechnet man dann eine Basis $\{c_1, \ldots, c_{n-r}\} \subset K^n$ für den Nullraum ab wie im Nullraumalgorithmus beschrieben. Dann fügt man die $n-r$ Zeilen $(c_i^H, 0, \ldots, 0) \in K^{m+n}$ unten an die Matrix $(\hat{A}, \hat{E}_m)$ an und bringt dann diese vergrößerte Matrix durch elementare Zeilenumformungen auf die in Schritt 2 genannte Gestalt

$$\tilde{\mathcal{A}} := \begin{pmatrix} E_n & B \\ 0 & D \end{pmatrix}.$$

Weiter geht es dann mit dem Schritt 3.

Beispiel 4. Zu berechnen ist die Pseudoinverse der Matrix

$$A := \begin{pmatrix} 14 & -4 & 10 & -6 \\ -4 & 5 & -5 & 6 \\ 10 & -5 & 11 & -4 \\ -6 & 6 & -4 & 10 \end{pmatrix}.$$

Man vergrößert jetzt A durch Anhängen der Einheitsmatrix E_4 zu $(A, E_4) \in K^{4\times 8}$. Gaußelimination führt dann auf die Zeilenstufenform

$$(\hat{A}, \hat{E}_4) = \left(\begin{array}{rrrr|rrrr} -4 & 5 & -5 & 6 & 0 & 1 & 0 & 0 \\ 0 & -3 & 7 & 2 & 0 & -3 & 0 & 2 \\ 0 & 0 & 16 & 16 & 0 & -5 & 1 & 5 \\ 0 & 0 & 0 & 0 & 2 & -5 & -3 & 3 \end{array}\right)$$

Hieraus liest man ab, dass der Nullraum von A eindimensional ist und der Vektor $(\frac{2}{3}, -\frac{5}{3}, -1, 1)^H$ diesen Raum aufspannt. Zur Vermeidung von Brüchen wird hier besser der Vektor $(-2, 5, 3, -3)^H$ genommen. Wir fügen daher $(-2, 5, 3, -3, 0, 0, 0, 0)$ als fünfte Zeile an die Matrix $(\hat{A}, \hat{E}_4)$ an,

$$\left(\begin{array}{rrrr|rrrr} -4 & 5 & -5 & 6 & 0 & 1 & 0 & 0 \\ 0 & -3 & 7 & 2 & 0 & -3 & 0 & 2 \\ 0 & 0 & 16 & 16 & 0 & -5 & 1 & 5 \\ 0 & 0 & 0 & 0 & 2 & -5 & -3 & 3 \\ -2 & 5 & 3 & -3 & 0 & 0 & 0 & 0 \end{array}\right),$$

und führen vollständige Gaußelimination für die ersten $n = 4$ Spalten durch,

$$\left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 0 & \frac{343}{752} & \frac{101}{752} & -\frac{175}{752} \\ 0 & 1 & 0 & 0 & 0 & \frac{247}{752} & \frac{53}{752} & -\frac{103}{752} \\ 0 & 0 & 1 & 0 & 0 & -\frac{209}{752} & \frac{13}{752} & \frac{145}{752} \\ 0 & 0 & 0 & 1 & 0 & -\frac{26}{752} & \frac{34}{752} & \frac{90}{752} \\ \hline 0 & 0 & 0 & 0 & 2 & -5 & -3 & 3 \end{array}\right).$$

Hieraus liest man

$$B = \frac{1}{752}\begin{pmatrix} 0 & 343 & 101 & -175 \\ 0 & 247 & 53 & -103 \\ 0 & -209 & 13 & 145 \\ 0 & -26 & 34 & 90 \end{pmatrix} \quad \text{und} \quad D = (2, -5, -3, 3)$$

ab. Mit $D^+ = (\frac{2}{47}, -\frac{5}{47}, -\frac{3}{47}, \frac{3}{47})^H$ gibt das schließlich

$$A^+ = B - BD^+D = \frac{1}{17672}\begin{pmatrix} 2543 & 1703 & -1441 & -298 \\ 1703 & 1574 & -1309 & 134 \\ -1441 & -1309 & 2467 & 1246 \\ -298 & 134 & 1246 & 1668 \end{pmatrix}.$$