Lösung 23: Sylvesters Trägheitssatz & Singulärwertzerlegung

1. Wir wissen, dass eine Basis $\mathcal{B}$ von $\mathbb{R}^n$ existiert, sodass

$$[\beta_Q]_{\mathcal{B}} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_{n-p-q} \end{pmatrix}$$

gilt. Da nach Annahme β_Q nicht ausgeartet ist, gilt $n = p + q$, da ansonsten für den eindeutigen Vektor $v \in V$ mit $[v]_{\mathcal{B}} = e_{p+q+1}$ folgt $\beta_Q(v, w) = 0$ für alle $w \in \mathbb{R}^n$, im Widerspruch zur Annahme.

Wir behaupten, dass $m = \min\{p, q\}$ die maximale Dimension eines vollständig isotropen Unterraumes ist und dass ein vollständig isotroper Unterraum der Dimension m existiert.

Zuerst zeigen wir, dass tatsächlich ein vollständig isotroper Unterraum der Dimension m existiert. Wir geben zuerst eine Skizze des Beweises um die Behauptung zu rechtfertigen. Im Anschluss liefern wir eine genaue Musterlösung. Seien $v, w \in \mathbb{R}^n$ die eindeutig bestimmten Elemente, sodass $[v]_{\mathcal{B}} = e_1$ sowie $[w]_{\mathcal{B}} = e_{p+1}$ gelten. Dann ist

$$\begin{aligned} \beta_Q(v+w, v+w) =& \beta_Q(v, v) + 2\beta_Q(v, w) + \beta_Q(w, w) \\ =& e_1^T e_1 + 2e_1^T e_{p+1} - e_{p+1}^T e_{p+1} = 0. \end{aligned}$$

Somit ist $\operatorname{span}(v + w)$ ein vollständig isotroper Unterraum von V. Wir definieren nun den Unterraum $U \subseteq \mathbb{R}^n$ durch $U = \operatorname{span}(v + w, v - w)$ und schreiben $\tilde{v} = v + w$ und $\tilde{w} = v - w$. Man beachte, dass $\beta_Q(\tilde{v}, \tilde{w}) = 2$ und insbesondere $\beta_Q|_{U\times U}$ nicht ausgeartet ist. Es gilt $\dim U = 2$, und jeder maximale, vollständig isotrope Unterraum von U ist eindimensional, denn jeder 0-dimensionale Unterraum ist gleich $\{0\}$ und somit in einem eindimensionalen, vollständig isotropen Unterraum – bspw. $\operatorname{span}(\tilde{v})$ – enthalten, und jeder zweidimensionale Unterraum von U ist gleich U und

Bitte wenden!

weil $\beta_Q|_{U\times U} \neq 0$ nicht vollständig isotrop. Da $\beta_Q|_{U\times U}$ nicht ausgeartet ist, folgt aus Aufgabe 4 in Serie 22, dass $\mathbb{R}^n = U \oplus U^{\beta_Q}$ gilt. Wir können zeigen, dass auf U^{β_Q} die Bilinearform β_Q die Darstellungsmatrix

$$\begin{pmatrix} I_{p-1} & \\ & -I_{q-1} \end{pmatrix}$$

besitzt. Nun machen wir auf dieselbe Art und Weise auf $\mathbb{R}^{n-2}$ zwei weiter und entfernen wieder ein Paar $\tilde{v}, \tilde{w}$ mit den entsprechenden Eigenschaften, bis entweder $p = 0$ oder $q = 0$ gilt. Auf dem verbliebenen Komplement ist die quadratische Form definit, und somit enthält dieser Unterraum keinen vollständig isotropen Unterraum mehr. Alle vorangegangenen, abgespaltenen Unterräume liefern je einen eindimensionalen, vollständig isotropen Unterraum und es folgt, dass ein vollständig isotroper Unterraum der Dimension $\min\{p, q\}$ existiert.

Wir verwenden Induktion über m. Wenn $p = 0$ (oder $q = 0$) ist, dann ist Q positiv (bzw. negativ) definit, und insbesondere ist kein nicht trivialer Unterraum vollständig isotrop. Insbesondere gilt die Behauptung im Falle $m = 0$. Angenommen, die Aussage stimmt für $m - 1$. Da der Fall $m = 0$ bereits bewiesen ist, können wir also annehmen, dass $pq \neq 0$. Wir wählen eine Basis $v_1, \ldots, v_p, w_1, \ldots, w_q$ von $\mathbb{R}^n$, sodass $\beta_Q(v_i, v_j) = \delta_{ij}$, $\beta_Q(w_k, w_l) = -\delta_{kl}$ und $\beta_Q(v_i, w_l) = 0$ für alle $1 \leq i, j \leq p$, $1 \leq k, l \leq q$. Wir zeigen, dass die Restriktion von β_Q auf den Unterraum

$$W_1 := \operatorname{span}\{v_1 + w_1, v_1 - w_1\}$$

nicht ausgeartet ist. Man berechnet

$$\begin{aligned}
\beta_Q(v_1 + w_1, v_1 + w_1) =& 0 \\
\beta_Q(v_1 - w_1, v_1 + w_1) =& \beta_Q(v_1 + w_1, v_1 - w_1) = 2 \\
\beta_Q(v_1 - w_1, v_1 - w_1) =& 0
\end{aligned}$$

und somit ist die Darstellungsmatrix von $\beta_Q|_{W_1\times W_1}$ bezüglich der geordneten Basis $\mathcal{B}_1 := (v_1 + w_1, v_1 - w_1)$ von W_1 invertierbar (mit Determinante -4). Insbesondere ist die Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $x \mapsto [\beta_Q|_{W_1\times W_1}]_{\mathcal{B}_1} x$ surjektiv. Angenommen es ist $w \in W_1 \setminus \{0\}$, dann ist $[w]_{\mathcal{B}_1} \neq 0$. Dann existiert ein $x \in \mathbb{R}^2$, sodass $[w]_{\mathcal{B}_1}^T [\beta_Q|_{W_1\times W_1}]_{\mathcal{B}_1} x \neq 0$ und insbesondere ein $w' \in W_1$, sodass

$$0 \neq [w]_{\mathcal{B}_1}^T [\beta_Q|_{W_1\times W_1}]_{\mathcal{B}_1} x = [w]_{\mathcal{B}_1}^T [\beta_Q|_{W_1\times W_1}]_{\mathcal{B}_1} [w']_{\mathcal{B}_1} = \beta_Q(w, w').$$

Aus Aufgabe 4, Serie 22, wissen wir, dass $\mathbb{R}^n = W_1^{\beta_Q} \oplus W_1$ gilt. Da

$$\mathcal{B}_1^* := (v_2, \ldots, v_p, w_2, \ldots, w_q)$$

eine Basis von $W_1^{\beta_Q}$ ist, besitzt $\beta_Q|_{W_1^{\beta_Q}\times W_1^{\beta_Q}}$ die Darstellungsmatrix

$$\begin{pmatrix} I_{p-1} & \\ & -I_{q-1} \end{pmatrix}.$$

Siehe nächstes Blatt!

Wir wissen, dass jeder vollständig isotrope Unterraum von $W_1^{\beta_Q}$ maximaler Dimension die Dimension $m-1$ hat. Sei $U \subseteq W_1^{\beta_Q}$ ein solcher Unterraum. Da $U \subseteq W_1^{\beta_Q}$ ist, ist $U + \operatorname{span}\{v_1 + w_1\}$ ein vollständig isotroper Unterraum von V. Seien nämlich $u \in U$ und $\lambda \in \mathbb{R}$, dann ist für $\tilde{v} := v_1 + w_1$

$$Q\big(u + \lambda\tilde{v}\big) = \beta_Q\big(u + \lambda\tilde{v}, u + \lambda\tilde{v}\big) = \beta_Q(u,u) + 2\lambda\beta_Q(u,\tilde{v}) + \lambda^2\beta_Q(\tilde{v},\tilde{v}) = 0,$$

da $u \in W_1^{\beta_Q}$ und nach Definition von U und $v_1 + w_1$. Somit ist $U + \mathbb{R}\tilde{v}$ ein vollständig isotroper Unterraum mit Dimension m. Dies beweist, dass $\mathbb{R}^n$ einen vollständig isotropen Unterraum der Dimension m besitzt.

Es verbleibt zu zeigen, dass m eine maximale Dimension ist. Wir definieren $V^+ := \operatorname{span}\{v_1, \dots, v_p\}$ und $V^- := \operatorname{span}\{w_1, \dots, w_q\}$. Sei $W \subseteq \mathbb{R}^n$ ein Unterraum von Dimension $r > p$. Dann gilt

$$\begin{aligned}\dim(W \cap V^-) &= \dim(W) + \dim(V^-) - \dim(W + V^-)\\ &\geq \dim(W) + \dim(V^-) - \dim(\mathbb{R}^n)\\ &= r + q - n > 0\end{aligned}$$

Insbesondere enthält $W \cap V^-$ ein von Null verschiedenes Element $v \in V^-$. Auf V^- ist aber β_Q nach Voraussetzung negativ definit und somit gilt $Q(v) < 0$. Somit ist W nicht vollständig isotrop. Dies zeigt, dass jeder vollständig isotrope Unterraum Dimension $r \leq p$ hat. Analog sei nun $W \subseteq \mathbb{R}^n$ ein Unterraum von Dimension $r > q$. Dann gilt

$$\begin{aligned}\dim(W \cap V^+) &= \dim(W) + \dim(V^+) - \dim(W + V^+)\\ &\geq \dim(W) + \dim(V^+) - \dim(\mathbb{R}^n)\\ &= r + q - n > 0\end{aligned}$$

Insbesondere enthält $W \cap V^+$ ein von Null verschiedenes Element $v \in V^+$. Auf V^+ ist aber β_Q nach Voraussetzung positiv definit und somit gilt $Q(v) < 0$. Somit ist W nicht vollständig isotrop. Dies zeigt, dass jeder vollständig isotrope Unterraum Dimension $r \leq q$ hat. Beides zusammen impliziert, dass jeder vollständig isotrope Unterraum Dimension $r \leq \min\{p, q\}$ besitzt.

2. **a)** Sei $A = RDQ$ mit $Q \in O(n)$ und $R \in O(m)$ eine Singulärwertzerlegung von A. Dann ist

$$AA^T = (RDQ)(Q^T D^T R^T) = RDD^T R^T$$

und somit DD^T ähnlich zu AA^T. Man beachte, dass

$$(DD^T)_{ij} = \sum_{k=1}^{m} D_{ik} D_{jk} = \delta_{ij} D_{ii} D_{jj}$$

Bitte wenden!

und somit ist DD^T eine Diagonalmatrix. Da DD^T ähnlich ist zu AA^T, sind die Diagonaleinträge genau die Eigenwerte von AA^T und somit durch AA^T eindeutig bestimmt. Da die Einträge von D alle nicht-negativ sind, sind die Wurzeln eindeutig bestimmt und somit die Einträge von D durch A vollständig bestimmt.

b) Wenn A symmetrisch positiv definit ist, dann ist A orthogonal diagonalisierbar, d.h. es gibt $Q \in O(n)$ und eine Diagonalmatrix D, sodass $QDQ^T = A$. Mittels Konjugation von D durch Elementarmatrizen vom Typ I können wir D in eine Diagonalmatrix D' überführen, deren Diagonaleinträge absteigend geordnet sind. Angenommen $D' = P^{-1}DP$ für P ein Produkt von Permutationsmatrizen (Elementarmatrizen vom Typ I). Weil für eine Elementarmatrix R vom Typ I die Gleichungen $RR = I_n$ und $R^T = R$ gelten, ist $R \in O(n)$ und somit $P \in O(n)$. Also ist $D = PD'D^T$ und folglich $A = QDQ^T = (QP)D'(QP)^T$. Da A positiv definit ist, sind die Diagonaleinträge von D' alle positiv, und somit ist $QD'Q^T = A$ eine Singulärwertzerlegung. Da die Diagonaleinträge von D' genau die Eigenwerte von A sind, stimmen die Singulärwerte mit den Eigenwerten überein.

c) Aus dem ersten Spektralsatz wissen wir, dass B orthogonal diagonalisierbar ist, mit positiven DIagonaleinträgen. Jede Diagonalmatrix mit positiven Diagonaleinträgen besitzt eine eindeutige diagonale, positiv definite Quadratwurzel, und somit ist $B = (Q^T\sqrt{D}Q)^2$ für ein $Q \in O(n)$.

Für die Eindeutigkeit seien also ganz allgemein A, B symmetrische, positiv definite Matrizen, sodass $A^2 = B$. Wir behaupten, dass wir eine Basis $(v_1, \dots, v_n)$ von V sowie positive Skalare $\lambda_1, \dots, \lambda_n$ finden können, sodass $Av_i = \lambda_i v_i$ gilt und v_i, λ_i durch B vollständig bestimmt sind. Dadurch ist A vollständig bestimmt, da die Matrizen $M = (v_1 \mid \cdots \mid v_n)$ sowie $\tilde{M} = (\lambda_1 v_1 \mid \cdots \mid \lambda_n v_n)$ invertierbar sind.

Da B symmetrisch, positiv definit ist, existiert eine ONB $(v_1, \dots, v_n)$ von V, bestehend aus Eigenvektoren von B zu strikt positiven Eigenwerten μ_i. Definiere $\lambda_i = \sqrt{\mu_i}$. Wir zeigen, dass v_i ein Eigenvektor von A zum Eigenwert λ_i ist. Sei $v = Av_i - \lambda_i v_i$, dann ist

$$Av = A^2 v_i - \lambda_i A v_i = \mu_i v_i - \lambda_i A v_i = -\lambda_i(Av_i - \lambda_i v_i) = -\lambda_i v$$

und wegen $\lambda_i > 0$ folgt $v = 0$. Ansonsten wäre v ein Eigenvektor von A zum Eigenwert $-\lambda_i < 0$, was unmöglich ist, da A positiv definit ist und somit alle Eigenwerte strikt positiv sind.

d) Da A symmetrisch, positiv definit ist, ist auch A^2 symmetrisch, positiv definit. Sei $A = RDQ^T$ eine Singulärwertzerlegung von A. Aus der Symmetrie folgt $A = A^T = QDR^T$ und somit $A^2 = (RDQ^T)^2$ und $A^2 = (QDQ^T)^2$. Aus vorangehender Teilaufgabe folgt, dass $RDQ^T = QDQ^T$, was wegen der Invertierbarkeit von Q und D impliziert, dass $Q = R$.

Siehe nächstes Blatt!

3. Im Folgenden seien $\mathcal{B} = (v_1, \ldots, v_n) \in V^n$ und $\mathcal{C} = (w_1, \ldots, w_m) \in W^m$ geordnete orthonormale Basen wie in der Singulärwertzerlegung von T, d.h. $Tv_i = \sigma_i w_i$ für alle $1 \leq i \leq r := \mathrm{Rang}(T)$. Wir bezeichnen mit λ_i die Eigenwerte von TT^* in absteigender Reihenfolge, d.h. $\lambda_i = \sigma_i^2$. Man berechnet

$$\|Tv\|_W^2 = \left\| \sum_{i=1}^r \sigma_i \langle v, v_i \rangle_V w_i \right\|_W^2 = \sum_{i=1}^r (\sigma_i \langle v, v_i \rangle_V)^2 \|w_i\|_W^2 = \sum_{i=1}^r \lambda_i \langle v, v_i \rangle_V^2.$$

Insbesondere ist

$$\frac{\|Tv\|_W^2}{\|v\|_V^2} = \frac{\sum_{i=1}^r \lambda_i \langle v, v_i \rangle_V^2}{\sum_{i=1}^r \langle v, v_i \rangle_V^2} \leq \lambda_1 \frac{\sum_{i=1}^r \langle v, v_i \rangle_V^2}{\sum_{i=1}^r \langle v, v_i \rangle_V^2} = \sigma_1^2$$

und das Maximum wird für den Vektor v_1 angenommen.

Andererseits ist

$$\frac{\|Tv\|_W^2}{\|v\|_V^2} = \frac{\sum_{i=1}^r \lambda_i \langle v, v_i \rangle_V^2}{\sum_{i=1}^r \langle v, v_i \rangle_V^2} \geq \lambda_r \frac{\sum_{i=1}^r \langle v, v_i \rangle_V^2}{\sum_{i=1}^r \langle v, v_i \rangle_V^2} = \sigma_r^2$$

und der Wert wird angenommen für $v = v_r$.

Wenn $r < n$, dann ist $\mathrm{Ker}(T) \neq 0$, und somit existiert ein $v \in V \setminus \{0\}$, sodass $\frac{\|Tv\|_W}{\|v\|_V} = 0$. Es folgt

$$\max_{v \in V \setminus \{0\}} \frac{\|Tv\|_W}{\|v\|_V} = \sigma_1 \text{ und } \min_{v \in V \setminus \{0\}} \frac{\|Tv\|_W}{\|v\|_V} = \begin{cases} \sigma_r & \text{falls } r = n \\ 0 & \text{sonst} \end{cases}$$

4. **a)** Im Folgenden schreiben wir $A = RDQ$ für eine Singulärwertzerlegung von A, wobei R, D und Q zu bestimmen sind. Wir berechnen die Matrix

$$AA^T = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Das zugehörige charakteristischen Polynom ist

$$\mathrm{char}_{AA^T}(X) = \det \begin{pmatrix} 2 - X & -1 \\ -1 & 2 - X \end{pmatrix} = X^2 - 4X + 3 = (X - 3)(X - 1).$$

Es folgt, dass die Matrix AA^T die Eigenwerte $\lambda_1 = 3$ und $\lambda_2 = 1$ besitzt. Die Singulärwerte von A sind also $\sigma_1 = \sqrt{3}$ und $\sigma_2 = 1$ und die Matrix D ist gegeben durch

$$D = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Bitte wenden!

Die Eigenvektoren von AA^T zu den Eigenwerten λ_1 bzw. λ_2 sind gegeben durch

$$u_1 = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ und } u_2 = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Die Spaltenvektoren von R sind gegeben durch die Eigenvektoren von AA^T, also erhalten wir

$$R = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Aus der Gleichung $R^T A = DQ$ folgt, dass für $i = 1, 2$ die i-te Zeile von Q das $\frac{1}{\sigma_i}$-fache der i-ten Zeile von $R^T A$ ist. Wir haben also

$$Q_{(1)} = \tfrac{1}{\sigma_1} u_1^T A = \tfrac{1}{\sqrt{6}}(2, -1, 1), \quad Q_{(2)} = \tfrac{1}{\sigma_2} u_2^T A = \tfrac{1}{\sqrt{2}}(0, 1, 1)$$

und einem beliebigen Vektor $Q_{(3)}$, sodass $(Q_{(1)}, Q_{(2)}, Q_{(3)})$ eine Orthonormalbasis ist. Durch Anwenden des Gram-Schmidt Algorithmus, zum Beispiel auf die Basis $(Q_{(1)}, Q_{(2)}, e_1)$ mit $e_1 = (1, 0, 0)$, wählen wir

$$\begin{aligned} Q_{(3)} &= \tfrac{1}{\|e_1 - \langle e_1, Q_{(1)}\rangle Q_{(1)} - \langle e_1, Q_{(2)}\rangle Q_{(2)}\|}(e_1 - \langle e_1, Q_{(1)}\rangle Q_{(1)} - \langle e_1, Q_{(2)}\rangle Q_{(2)}) \\ &= \tfrac{1}{\sqrt{3}}(1, 1, -1) \end{aligned}$$

Es folgt $A = RDQ$, wobei

$$Q = \tfrac{1}{\sqrt{6}} \begin{pmatrix} 2 & -1 & 1 \\ 0 & \sqrt{3} & \sqrt{3} \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} \end{pmatrix}.$$

b) Die Transponierte einer orthogonalen Matrix ist wieder orthogonal. Somit ist $A^T = Q^T D^T R^T$ eine mögliche gesuchte Zerlegung.

5. **a)** Wir wissen aus Serie 19, Aufgabe 5, dass $\mathrm{Rang}(A^T A) = \mathrm{Rang}(A)$ und somit hat $A^T A$ tatsächlich vollen Rang und ist invertierbar. Im Folgenden seien $K \in \mathrm{SO}(m)$, $L \in \mathrm{SO}(n)$ und $D \in M_{m\times n}(\mathbb{R})$ so gewählt, dass $A = KDL$ eine Singulärwertzerlegung von A ist. Wir wissen aus dem vergangenen Semester, dass $\mathrm{Rang}(A) = \mathrm{Rang}(D)$ und somit existiert eine invertierbare Diagonalmatrix $E \in \mathrm{GL}_n(\mathbb{R})$, sodass

$$D = \left(\frac{E}{\mathbf{0}_{(m-n)\times n}} \right).$$

Man berechnet

$$\begin{aligned} A(A^T A)^{-1} A^T &= KDL(L^T D^T K^T KDL)^{-1} L^T D^T K^T \\ &= KDL(L^T E^2 K)^{-1} L^T D^T K^T \end{aligned}$$

Siehe nächstes Blatt!

$$
\begin{aligned}
&= KDE^{-2}D^TK^T \\
&= K \begin{pmatrix} E \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix} E^{-2} \begin{pmatrix} E & \mathbf{0}_{n\times(m-n)} \end{pmatrix} K^T \\
&= K \begin{pmatrix} I_n \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix} \begin{pmatrix} I_n & \mathbf{0}_{n\times(m-n)} \end{pmatrix} K^T \\
&= K \begin{pmatrix} I_n & \mathbf{0}_{n\times(m-n)} \\ \mathbf{0}_{(m-n)\times n} & \mathbf{0}_{(m-n)\times(m-n)} \end{pmatrix} K^T
\end{aligned}
$$

Sei nun $\mathcal{C} = (w_1, \dots, w_m)$ die Orthonormalbasis von $\mathbb{R}^m$ zur Singulärwertzerlegung von A, dann ist also $K^{(i)} = w_i$ und somit folgt

$$A(A^TA)^{-1}A^Tw_i = K \begin{pmatrix} I_n & \mathbf{0}_{n\times(m-n)} \\ \mathbf{0}_{(m-n)\times n} & \mathbf{0}_{(m-n)\times(m-n)} \end{pmatrix} e_i = \begin{cases} w_i & \text{falls } i \leq n \\ 0 & \text{sonst} \end{cases}$$

Da nach Voraussetzung bzw. nach Definition der Singulärwertzerlegung die geordnete Menge $(w_1, \dots, w_n)$ eine Basis von $\mathrm{Im}(A)$ ist, stimmen die Abbildungen P_A und $L_{A(A^TA)^{-1}A}$ auf einer Basis von $\mathbb{R}^m$ überein. Aus der Linearität folgt $P_A = L_{A(A^TA)^{-1}A}$ und somit auch

$$[P_A]_{\mathcal{E}_m} = [L_{A(A^TA)^{-1}A}]_{\mathcal{E}_m} = A(A^TA)^{-1}A.$$

b) Sei A die Matrix mit Spalten $A^{(1)} = \mathbf{1}$ und $A^{(j)} = x^{(j-1)}$ für $j > 1$. Dann ist nach Voraussetzung $A \in M_{m\times(n+1)}(\mathbb{R})$ mit $\mathrm{Rang}(A) = n+1$ und somit

$$\min\{\|y - w\|_2 | w \in \mathrm{Im}(A)\} = \|y - P_Ay\|_2 = \|y - A(A^TA)^{-1}A^Ty\|_2.$$

Es folgt

$$\min_{\beta\in\mathrm{R}^{n+1}} \|y - A\beta\|_2 = \|y - A(A^TA)^{-1}Ay\|_2$$

und da die orthogonale Projektion auf einen Unterraum eindeutig durch die minimierende Eigenschaft bestimmt ist, folgt

$$A\hat{\beta} = A(A^TA)^{-1}Ay.$$

Da $\mathrm{Rang}(A) = n+1$, folgt $\mathrm{Ker}(A) = \{0\}$, d.h. L_A ist injektiv und somit folgt $\hat{\beta} = (A^TA)^{-1}Ay$.