

# Formale Grundlagen

## 4. Übungsaufgaben

2008-04-22, Lösungen

### 1. Wir suchen die beste Lösung der Gleichung

$$3x_1 + 1x_2 = 6.$$

Schreiben Sie diese Gleichung mit Matrizen und Vektoren, also in der Form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}.$$

*Lösung:*

$$(3 \ 1) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (6)$$

Dabei bezeichne  $\mathbf{x}$  den Spaltenvektor  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ . In diesem einfachen Beispiel hat  $\mathbf{A}$  nur eine Zeile,  $\mathbf{A} = (3 \ 1)$ , und  $\mathbf{b}$  ist eine  $1 \times 1$ -Matrix. Zeichnen Sie den Vektor  $\mathbf{z}_1$ , der die einzige Zeile von  $\mathbf{A}$  beschreibt, in ein kartesisches Koordinatensystem.

Wir lösen zuerst die homogene Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{o}$ . Man findet leicht ein paar Lösungen, z.B. indem man  $x_1 = 0$  setzt und dann nach  $x_2$  auflöst, oder indem man  $x_1 = 1$  setzt. Die Gesamtheit der Lösungen (der Nullraum von  $\mathbf{A}$ ) ist jedenfalls ein linearer Unterraum. Finden Sie dafür eine Basis und zeichnen Sie alles in das kartesische Koordinatensystem von vorhin. Berechnen Sie das Skalarprodukt von jeder Zeile von  $\mathbf{A}$  mit jedem Basisvektor des Nullraums. Schließen Sie daraus, daß der Nullraum orthogonal zu den Zeilen von  $\mathbf{A}$  ist. Überprüfen Sie dies auch anhand der Skizze.

*Lösung:*

$x_1 = 0$  führt zur trivialen Lösung  $x_1 = x_2 = 0$ . Setzen wir dagegen  $x_1 = -1$ , so ergibt sich  $x_2 = 3$ . Damit ist

$$\mathbf{n}_1 := \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

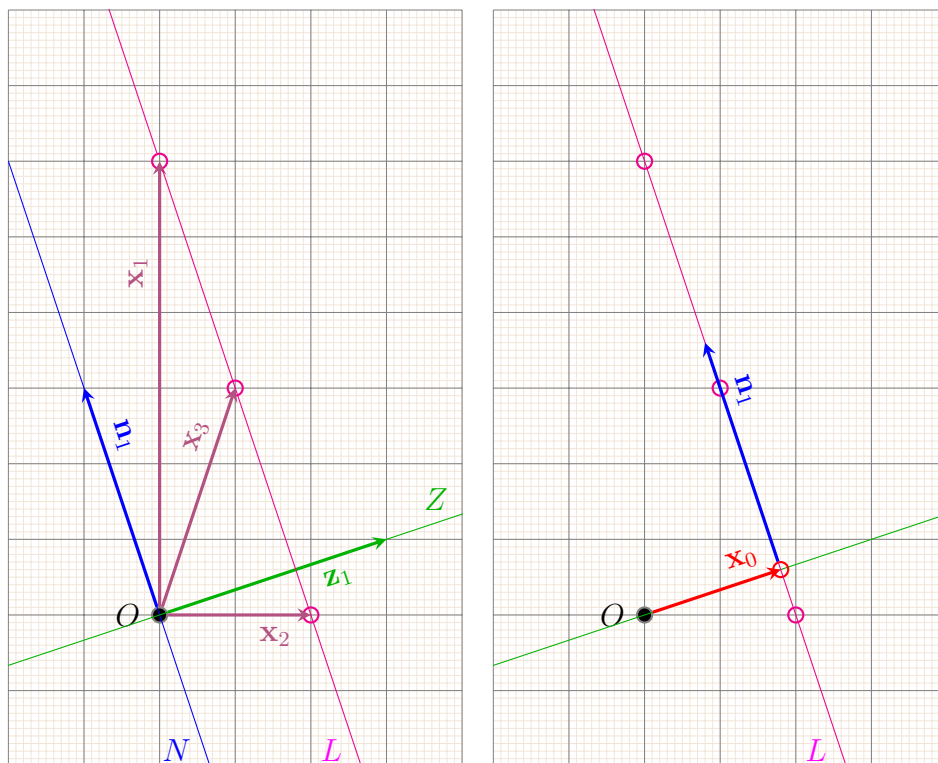


Abbildung 1: Zeichnungen zu Beispiel ??

eine nichttriviale Lösung. Alle anderen Lösungen sind Vielfache davon. Der Nullraum  $N$  ist somit gegeben durch  $N = L(\mathbf{n}_1)$ . Aus der Zeichnung erkennen wir, daß  $N$  orthogonal zu  $\mathbf{z}_1$  erscheint. Tatsächlich rechnet man leicht nach, daß  $\langle \mathbf{z}_1, \mathbf{n}_1 \rangle = 0$ . Wir bemerken, daß dies kein Zufall ist, sondern genau der Gleichung entspricht. Wir erkennen somit, daß der Nullraum orthogonal zum Zeilenraum  $Z$  ist, wobei  $Z = L(\mathbf{z}_1)$ .

Nun suchen wir die Lösungen der inhomogenen Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ . Wir können wieder, z.B. durch Nullsetzen einer Variablen, leicht Lösungen  $\mathbf{x}_1, \mathbf{x}_2$  finden. Die Gesamtheit der Lösungen ist nun ein affiner Unterraum  $L$ .

*Lösung:* Aus  $x_1 = 0$  erhalten wir  $x_2 = 6$ . Somit ist

$$\mathbf{x}_1 = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$

eine erste Lösung. Aus  $x_2 = 0$  erhalten wir  $x_1 = 2$ , somit ist

$$\mathbf{x}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

eine weitere Lösung. Da auch alle Affinkombinationen von Lösungen wieder Lösungen sind, verbinden wir diese beiden Punkte durch eine Gerade. Dieser affine Unterraum  $L$  ist parallel zum Nullraum  $N$  und enthält alle Lösungen. Setzen wir z.B.  $x_1 = 1$ , so erhalten wir eine dritte Lösung

$$\mathbf{x}_3 = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

welche tatsächlich wieder in  $L$  liegt.

Wir suchen nun unter allen Lösungen die kleinste, also eine Lösung  $\mathbf{x}_0$  sodaß  $\|\mathbf{x}_0\| \leq \|\mathbf{x}\|$ , für jede Lösung  $\mathbf{x}$  von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ . In der Skizze ist diese Lösung leicht zu finden, da geometrisch klar ist, daß  $\mathbf{x}_0$  orthogonal zu  $L$  sein muß. Tatsächlich ist nur der Vektor  $\mathbf{z}_1$  entsprechend zu skalieren. Ist das Zufall? Der gesuchte Vektor hat also die Form  $\mathbf{z}_1 \cdot y_1$ , wobei  $y_1$  der passende Skalar ist. Man beachtet, daß  $\mathbf{z}_1$  zwar die Zeile von  $\mathbf{A}$  beschreibt, die Lösung  $\mathbf{z}_1 \cdot y_1$  aber ein Spaltenvektor sein sollte. Wir bezeichnen sie daher mit  $\mathbf{A}^T \cdot \mathbf{y}$ , wobei  $\mathbf{A}^T$  die transponierte Matrix von  $\mathbf{A}$  ist und  $\mathbf{y} = (y_1)$ .

*Lösung:* Da der gesuchte Vektor  $\mathbf{x}_0$  ein Vielfaches von  $\mathbf{z}_1$  sein soll, verwenden wir den Ansatz

$$\mathbf{x}_0 = \mathbf{z}_1 \cdot y_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot y_1 = \mathbf{A}^T \cdot \mathbf{y}.$$

Wir setzen also  $\mathbf{A}^T \cdot \mathbf{y}$  für  $\mathbf{x}$  ein und haben insgesamt die Gleichung  $\mathbf{A} \cdot \mathbf{A}^T \cdot \mathbf{y} = \mathbf{b}$  nach  $\mathbf{y}$  zu lösen.

*Lösung:* Durch Einsetzen erhalten wir somit die Gleichung

$$\begin{pmatrix} 3 & 1 \end{pmatrix} \cdot \left( \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot (y_1) \right) = (6)$$

Umklammern ergibt dann

$$\left( \begin{pmatrix} 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right) \cdot (y_1) = (6),$$

was ausgerechnet zu

$$10 \cdot (y_1) = (6),$$

führt, sodaß wir als Lösung

$$y_1 = \frac{6}{10}$$

erhalten. Dies setzen wir schließlich in den Ansatz für  $\mathbf{x}$  ein, und erhalten

$$\mathbf{x}_0 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot \frac{6}{10} = \begin{pmatrix} \frac{18}{10} \\ \frac{6}{10} \end{pmatrix}$$

als optimale Lösung.

Wir können diese sogar als  $\mathbf{y} = (\mathbf{A} \cdot \mathbf{A}^T)^{-1} \cdot \mathbf{b}$  darstellen. Insgesamt ergibt sich somit  $\mathbf{x} = \mathbf{A}^T \cdot (\mathbf{A} \cdot \mathbf{A}^T)^{-1} \cdot \mathbf{b}$ . Die Matrix  $\mathbf{A}^T \cdot (\mathbf{A} \cdot \mathbf{A}^T)^{-1}$  nennen wir Moore-Penrose-Inverse und bezeichnen sie mit  $\mathbf{A}^-$ . Damit erhalten wir als (optimale) Lösung der Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  einfach  $\mathbf{x} = \mathbf{A}^- \cdot \mathbf{b}$ . Vergewissern Sie sich, daß das errechnete Ergebnis mit der Skizze übereinstimmt.

## 2. Wir lösen nun das Gleichungssystem

$$\begin{aligned} 3y_1 &= 2 \\ 1y_1 &= 0. \end{aligned}$$

Schreiben Sie dieses in Matrixform  $\mathbf{B} \cdot \mathbf{y} = \mathbf{c}$ . Es läßt sich auch interpretieren als das Problem, für den Vektor  $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$  die Koordinaten bezüglich der Basis  $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$  zu finden. Machen Sie davon wieder eine Skizze.

*Lösung:* Das Gleichungssystem läßt sich schreiben als

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot (y_1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Also ist

$$\mathbf{B} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \mathbf{y} = (y_1) \quad \mathbf{c} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Offensichtlich ist  $\mathbf{c}$  kein Vielfaches von  $\mathbf{s}_1$ , der (einzigen) Spalte von  $\mathbf{B}$ , und daher nicht in der linearen Hülle der Spalten von  $\mathbf{A}$ .

Klarerweise ist dieses Problem unlösbar. Wir suchen daher eine Lösung, die den Fehler minimiert, d.h. ein  $\mathbf{y}_0$ , sodaß  $\|\mathbf{B} \cdot \mathbf{y}_0 - \mathbf{c}\| \leq \|\mathbf{B} \cdot \mathbf{y} - \mathbf{c}\|$ , für alle  $\mathbf{y}$ .

*Lösung:*  $\mathbf{B} \cdot \mathbf{y}_0$  sollte somit jener Punkt der Geraden  $S$  sein, welcher am nächsten beim Punkt  $\mathbf{c}$  liegt. Dieser entspricht offenbar der orthogonalen Projektion, d.h. der Vektor  $\mathbf{n} := \mathbf{c} - \mathbf{B} \cdot \mathbf{y}_0$  steht im rechten Winkel auf  $\mathbf{s}_1$ , also

$$\langle \mathbf{s}_1, \mathbf{n} \rangle = 0$$

was sich auch als

$$\mathbf{B}^T \cdot (\mathbf{c} - \mathbf{B} \cdot \mathbf{y}_0) = 0$$

schreiben läßt. Ausmultiplizieren ergibt dann

$$\mathbf{B}^T \cdot \mathbf{B} \cdot \mathbf{y}_0 = \mathbf{B}^T \cdot \mathbf{c}$$

bzw. konkret

$$(3 \ 1) \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \cdot (y_1) = (3 \ 1) \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

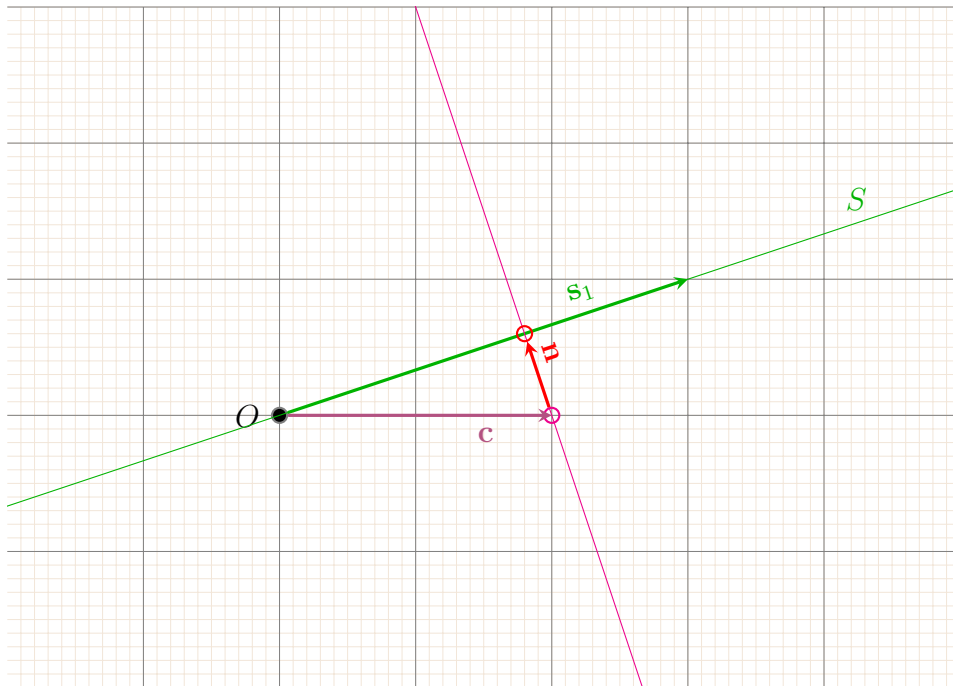


Abbildung 2: Zeichnung zu Beispiel ??

Dies ist leicht zu vereinfachen zu

$$10 \cdot y_1 = 6,$$

und somit ergibt sich die Lösung

$$y_1 = \frac{6}{10}.$$

Die Koordinaten des projizierten Vektors sind somit

$$\mathbf{B} \cdot y_1 = (3 \ 1) \cdot \left(\frac{6}{10}\right) = \begin{pmatrix} \frac{18}{10} \\ \frac{6}{10} \end{pmatrix}$$

Es stellt sich heraus, daß man diese Lösung bekommt, indem man die Gleichung auf beiden Seiten von links mit  $\mathbf{B}^T$  multipliziert. (Können Sie diesen Schritt exakt begründen?) Lösen Sie diese Gleichung und verifizieren Sie das Ergebnis anhand der Skizze. Die Matrix  $(\mathbf{B}^T \cdot \mathbf{B})^{-1} \cdot \mathbf{B}^T$  nennen wir wieder Moore-Penrose-Inverse und bezeichnen sie mit  $\mathbf{B}^-$ , sodaß wir  $\mathbf{y} = \mathbf{B}^- \cdot \mathbf{c}$  als bestapproximierende Lösung der (unlösbaren) Gleichung  $\mathbf{B} \cdot \mathbf{y} = \mathbf{c}$  erhalten.

3. Wenn die Matrix  $\mathbf{A} \cdot \mathbf{A}^T$  invertierbar ist, definieren wir

$$\mathbf{A}^- := \mathbf{A}^T \cdot (\mathbf{A} \cdot \mathbf{A}^T)^{-1},$$

und wenn die Matrix  $\mathbf{A}^T \cdot \mathbf{A}$  invertierbar ist, definieren wir

$$\mathbf{A}^- := (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T.$$

Zeigen Sie, daß in jedem Fall die Beziehungen

$$\mathbf{A} \cdot \mathbf{A}^- \cdot \mathbf{A} = \mathbf{A}^- \quad \mathbf{A}^- \cdot \mathbf{A}^- \cdot \mathbf{A}^- = \mathbf{A}^-$$

gelten.

*Lösung:* Im ersten Fall gilt

$$\mathbf{A} \cdot \mathbf{A}^- := \mathbf{A} \cdot \mathbf{A}^T \cdot (\mathbf{A} \cdot \mathbf{A}^T)^{-1} = \mathbf{E},$$

während im zweiten Fall

$$\mathbf{A}^- \cdot \mathbf{A} := (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{A} = \mathbf{E}.$$

Man nennt daher  $\mathbf{A}$  in diesen Fällen *einseitig invertierbar*. Dadurch fällt in jedem der Ausdrücke  $\mathbf{A} \cdot \mathbf{A}^- \cdot \mathbf{A}$  und  $\mathbf{A}^- \cdot \mathbf{A}^- \cdot \mathbf{A}^-$  stets der linke bzw. rechte Teil weg.

Es läßt sich ferner zeigen, daß es zu jeder Matrix  $\mathbf{A}$  genau eine Matrix  $\mathbf{A}^-$  gibt, sodaß die beiden obigen Gleichungen erfüllt sind und  $\mathbf{A} \cdot \mathbf{A}^-$  und  $\mathbf{A}^- \cdot \mathbf{A}$  symmetrisch sind. Sie heißt *Moore-Penrose-Inverse*. Wenn Sie gerade kein passendes Programm zur Verfügung haben, um diese zu berechnen, hilft wahrscheinlich ein Online-Matrizenrechner,

z.B. <http://www.bluebit.gr/matrix-calculator>

Berechnen Sie damit die Moore-Penrose-Inverse von ein paar Matrizen, insbesondere von denen in den vorigen Beispielen.

*Lösung:*

$$\begin{pmatrix} 3 & 1 \end{pmatrix}^- = \begin{pmatrix} \frac{3}{10} \\ \frac{1}{10} \end{pmatrix} \quad \begin{pmatrix} 3 \\ 1 \end{pmatrix}^- = \begin{pmatrix} \frac{3}{10} & \frac{1}{10} \end{pmatrix}$$

4. Es seien Punkte gegeben durch die Koordinaten  $(x_k, y_k)$  gemäß der Tabelle

$k$	$x_k$	$y_k$
1	0	2.5
2	1	3.7
3	3	5.5
4	7	10.9
5	8.3	12.2

Zeichnen Sie diese in ein kartesisches Koordinatensystem. Wir haben die Vermutung, daß diese alle auf einer Geraden der Form  $y = k \cdot x + d$  liegen, d.h. wir möchten alle Gleichungen  $y_k = k \cdot x_k + d$ , für  $k = 1, \dots, 5$  zu erfüllen. Klarerweise ist dies unlösbar, sodaß wir versuchen eine Lösung zu finden, die wenigstens den Fehler minimiert. Schreiben Sie daher das Gleichungssystem in der Form  $\mathbf{B} \cdot \mathbf{y} = \mathbf{c}$ , mit passenden Matrizen bzw. Vektoren, an. Beachten Sie, daß hier  $k$  und  $d$  die Variablen sind (also  $\mathbf{y} = \begin{pmatrix} d \\ k \end{pmatrix}$ ), während die  $x_k$  und  $y_k$  ja bekannt sind.

*Hinweis:* Sie können die Lösung mittels der Moore-Penrose-Inversen  $\mathbf{B}^-$  berechnen, so wie in den Beispielen zuvor.

*Lösung:* Die Gerade sollte idealerweise die folgenden Gleichungen erfüllen

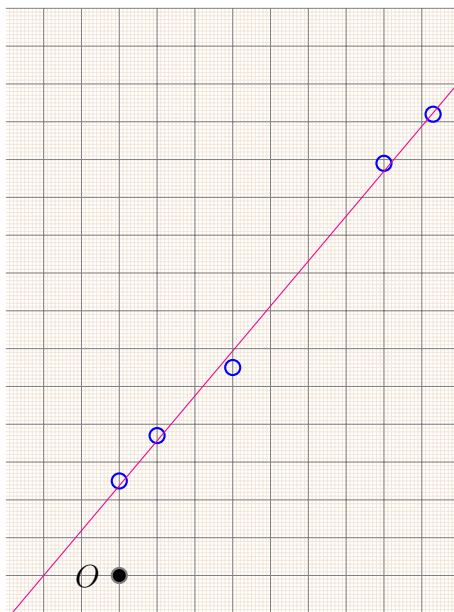


Abbildung 3: Ausgleichsgerade

$$k \cdot 0 + d = 2.5$$

$$k \cdot 1 + d = 3.7$$

$$k \cdot 3 + d = 5.5$$

$$k \cdot 7 + d = 10.9$$

$$k \cdot 8.3 + d = 12.2$$

Wir schreiben dies in Standardform um:

$$\begin{aligned}0 \cdot k + 1 \cdot d &= 2.5 \\1 \cdot k + 1 \cdot d &= 3.7 \\3 \cdot k + 1 \cdot d &= 5.5 \\7 \cdot k + 1 \cdot d &= 10.9 \\8.3 \cdot k + 1 \cdot d &= 12.2\end{aligned}$$

wodurch sich die Matrizenform ergibt als

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 7 & 1 \\ 8.3 & 1 \end{pmatrix} \cdot \begin{pmatrix} k \\ d \end{pmatrix} = \begin{pmatrix} 2.5 \\ 3.7 \\ 5.5 \\ 10.9 \\ 12.2 \end{pmatrix}$$

Da diese Punkte nicht auf einer Geraden liegen, ist dieses Gleichungssystem unlösbar. Wir suchen daher eine Ausgleichsgerade, welche die Abstände von den gegebenen Punkten minimiert.

Die Lösung ist somit einfach

$$\begin{pmatrix} k \\ d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 7 & 1 \\ 8.3 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 2.5 \\ 3.7 \\ 5.5 \\ 10.9 \\ 12.2 \end{pmatrix}$$

Und Ausrechnen ergibt

$$\begin{pmatrix} k \\ d \end{pmatrix} \approx \begin{pmatrix} 1.18804 \\ 2.37415 \end{pmatrix},$$

also die Gerade

$$y = 1.18804x + 2.37415$$

5. Gegeben sei eine Ebene  $E$  im  $\mathbb{R}^3$ , aufgespannt durch zwei Basisvektoren und ein weiterer Vektor:

$$E = L\left(\begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \\ -1 \end{pmatrix}\right) \quad \mathbf{v} = \begin{pmatrix} 3 \\ 2 \\ 8 \end{pmatrix}$$

Bestimmen Sie jenen Punkt der Ebene, welche am nächsten bei  $\mathbf{v}$  liegt (die Projektion von  $\mathbf{v}$  in die Ebene  $E$ ).

*Hinweis:* Versuchen Sie die Koordinaten von  $\mathbf{v}$  bezüglich den beiden Basisvektoren von  $E$  zu bestimmen. Die führt zu einem überbestimmten Gleichungssystem (3 Gleichungen und zwei Variable). Lösen Sie dieses wieder mit der Moore-Penrose-Inversen.

Auch hier ist es sinnvoll, eine Skizze anzufertigen.

*Lösung:*

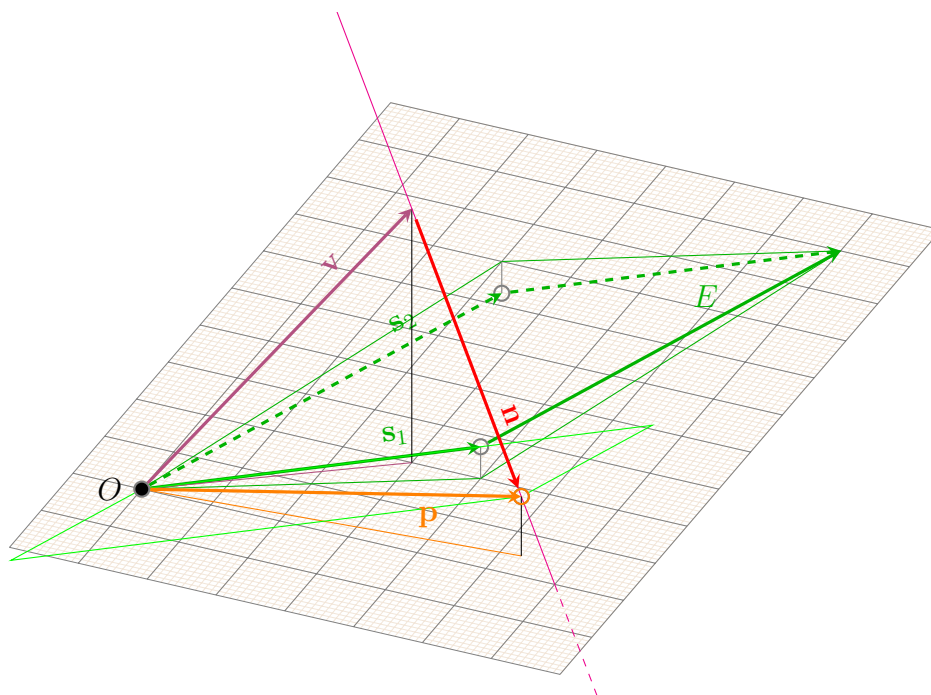


Abbildung 4: Zeichnung zu Beispiel ??

Der Vektor  $\mathbf{v}$  liegt nicht in der Ebene  $E$  (in der Zeichnung durch das dunkelgrüne Parallelogramm angedeutet). Von allen Vektoren dieser Ebene kommt der Vektor  $\mathbf{p}$ , die Projektion von  $\mathbf{v}$  in die Ebene, dem Vektor  $\mathbf{v}$  am nächsten. Die Projektionsrichtung  $\mathbf{n} := \mathbf{p} - \mathbf{v}$  hat die Eigenschaft, daß sie zu jedem Vektor der Ebene orthogonal ist, insbesondere auf  $\mathbf{s}_1$  und  $\mathbf{s}_2$ . Es muß also gelten:

$$\langle \mathbf{s}_1, \mathbf{n} \rangle = 0 \quad \langle \mathbf{s}_2, \mathbf{n} \rangle = 0,$$

was wir auch als

$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \cdot \mathbf{n} = \mathbf{o}$$

schreiben können. Mit der Matrix  $\mathbf{B} = (\mathbf{s}_1 \ \mathbf{s}_2)$  ist das einfach die Gleichung  $\mathbf{B}^T \cdot \mathbf{n} = \mathbf{o}$ . Da  $\mathbf{p}$  in der Ebene liegt, hat es die Form  $\mathbf{p} = \mathbf{B} \cdot \mathbf{y}$ , sodaß wir

erhalten

$$\mathbf{B}^T \cdot (\mathbf{B} \cdot \mathbf{y} - \mathbf{v}) = \mathbf{o}.$$

Einfaches Umformen führt somit zur Gleichung  $\mathbf{B}^T \cdot \mathbf{B} \cdot \mathbf{y} = \mathbf{B} \cdot \mathbf{v}$ , welche sich als  $\mathbf{y} = (\mathbf{B}^T \cdot \mathbf{B})^{-1} \cdot \mathbf{B}^T \cdot \mathbf{v}$  lösen läßt, falls  $\mathbf{B}^T \cdot \mathbf{B}$  invertierbar ist. In jedem Fall (auch wenn diese Invertierbarkeit nicht erfüllt ist) erhalten wir

$$\mathbf{y} = \mathbf{B}^- \cdot \mathbf{v}$$

bzw. hier konkret

$$\mathbf{y} = \begin{pmatrix} 4 & 2 \\ 2 & 7 \\ 1 & -1 \end{pmatrix}^- \cdot \begin{pmatrix} 3 \\ 2 \\ 8 \end{pmatrix} \approx \begin{pmatrix} 1.50649 \\ -0.363636 \end{pmatrix}$$

Dies sind die Koordinaten des projizierten Vektors  $\mathbf{p}$  bezüglich den Basisvektoren  $\mathbf{s}_1$  und  $\mathbf{s}_2$  der Ebene  $E$  (vgl. hellgrünes Parallelogramm; dies liegt in  $E$ ).

Den projizierten Vektor selbst erhalten wir somit durch

$$\mathbf{p} = \mathbf{B} \cdot \mathbf{y} \approx \begin{pmatrix} 5.2987 \\ 0.46753 \\ 1.87013 \end{pmatrix}$$