1) Organisatorisches

- Konzept EA
- Lösungshinweise (Gern im Forum nach Tipps fragen!)
- Klausurzulassung (Keine Beschränkung! Bearbeitung EA freiwillig)
- Lösungen:
  - → .pdf der EA werden bereitgestellt
  - → handschriftl. Lösungen der "Übungsaufgaben" aus dem Skript werden hochgeladen
- Plan für Online-Übungen: Lösungen zu EA / Fragen besprechen, ggf. eigene Lösungen vorstellen
- Forumsfragen

2) Beispiel Singulärwertzerlegung (Eigenwerte, Gram-Schmidt, etc.)

3) .. Lösungen / Tipps EA 1

Beispiel: Singulärwertzerlegung einer Matrix bestimmen

$$A = \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix} \qquad A \overset{!}{=} U \Sigma V^T$$

0. Suche im Skript nach relevanten Informationen / Stichworten

Bemerkung 1.3.7. Man beachte, dass der obige Beweis auch eine Möglichkeit zeigt, die Matrizen U, V und Σ zu konstruieren. Konkret kann man dies wie folgt tun:

1. Man bestimme die Eigenwerte $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ von $A^T A$ und berechne eine Orthonormalbasis $v_1, \ldots, v_n$ aus Eigenvektoren. Setze $V := (v_1, \cdots, v_n)$.
2. Man definiere $u_1, \ldots, u_k$ durch $u_i = Av_i / \|Av_i\|$ und nutze das Gram-Schmidtsche Orthogonalisierungsverfahren, um eine Orthonormalbasis $u_1, \ldots, u_m$ zu konstruieren. Setze $U := (u_1, \cdots, u_m)$.
3. Definiere Σ wie in (1.1).

Weiter definieren wir $\Sigma = (\sigma_{ij}) \in \mathbb{R}^{m \times n}$ durch

$$\sigma_{ij} := \begin{cases} \sqrt{\lambda_i} & \text{für } i = j \text{ und } i \leq k, \\ 0 & \text{sonst.} \end{cases} \tag{1.1}$$

---

1. Matrix V, bzw. V^T

a) $A^T A$ berechnen

$(-4)(-4) + (\frac{21}{5})(\frac{21}{5}) + 0 \cdot 0$

$$\begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix}^T \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -4 & \frac{21}{5} & 0 \\ -3 & -\frac{28}{5} & 0 \end{pmatrix} \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{841}{25} & -\frac{288}{25} \\ -\frac{288}{25} & \frac{1009}{25} \end{pmatrix}$$

b) Eigenwerte von $A^T A$

Definition 1.2.6. Sei $A \in \mathbb{K}^{n \times n}$ und $I \in \mathbb{K}^{n \times n}$ die Einheitsmatrix. Dann heißt das Polynom

$$p_A(z) := \det(A - zI)$$

in der Variablen $z \in \mathbb{K}$ *charakteristisches Polynom* von A.

Übungsaufgabe 1.2.7. *Man zeige, dass $p_A(z)$ tatsächlich ein Polynom vom Grad n ist.*

Satz 1.2.8. *Sei $A \in \mathbb{K}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) A besitzt den Eigenwert $\lambda \in \mathbb{K}$,

(ii) λ ist Nullstelle des charakteristischen Polynoms, d.h. $p_A(\lambda) = 0$.

$$p_{A^TA}(z) = \det(A^T A - z \cdot I)$$

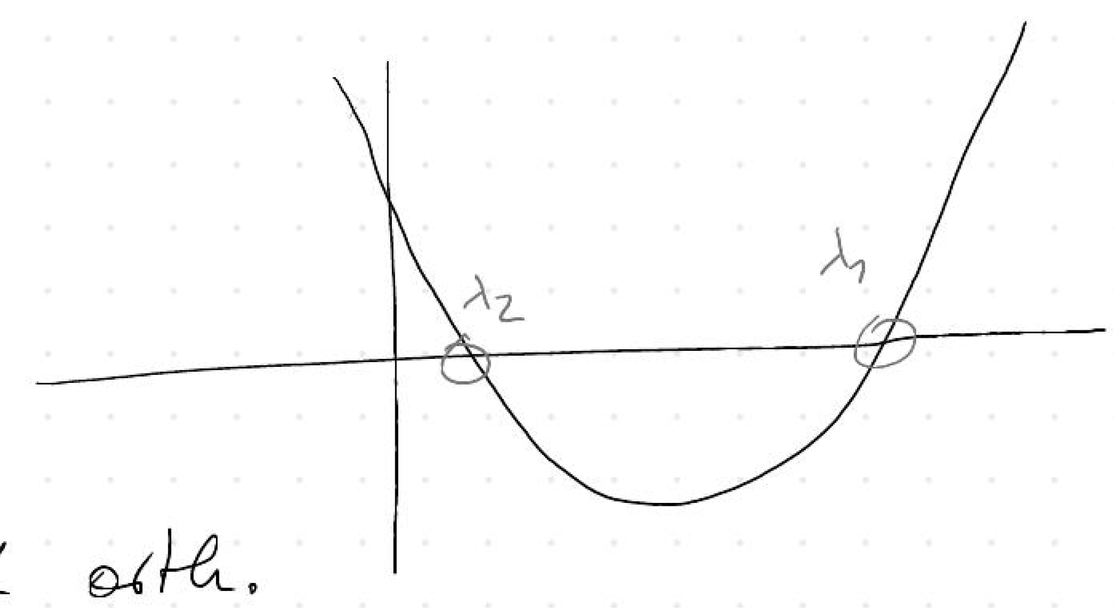
$$= \det \begin{pmatrix} \frac{841}{25} - z & -\frac{288}{25} \\ -\frac{288}{25} & \frac{1009}{25} - z \end{pmatrix}$$

$$= \left(\frac{841}{25} - z\right)\left(\frac{1009}{25} - z\right) - \left(-\frac{288}{25}\right)^2$$

$$= z^2 - 74z + 1225$$

$p_{A^TA}(\lambda) \overset{!}{=} 0 \quad \Rightarrow \quad \lambda^2 - 74\lambda + 1225 = 0$

p-q-Formel $\Rightarrow \quad \lambda_1 = 49 \;, \quad \lambda_2 = 25$



c) ONB aus Eigenvektoren

- A^TA schon symmetrisch $\Rightarrow$ alle Eigenräume sind orth. zueinander
- Die Eigenräume hier haben höchstens (und mindestens) Dimension 1
- Wir berechnen hier also jeweils einen EV & normieren diesen dann
- Falls ein Eigenraum eine höhere Dimension hätte, könnten wir das Gram-Schmidt-Verfahren anwenden (→ siehe unten)

Definition 1.2.1. Sei $A \in \mathbb{K}^{n\times n}$. Gibt es eine Zahl $\lambda \in \mathbb{K}$ und einen Vektor $v \in \mathbb{K}^n$, $v \neq 0$, mit der Eigenschaft

$$Av = \lambda v,$$

so heißt λ *Eigenwert* zum *Eigenvektor* v. Die Menge

$A^TA\, v = \lambda_1 v \iff (A^TA - \lambda_1 I)v = 0 \quad \Rightarrow$ Gleichungssysteme lösen

$$\begin{pmatrix} \frac{841}{25} - 49 & -\frac{288}{25} \\ -\frac{288}{25} & \frac{1009}{25} - 49 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -\frac{384}{25} & -\frac{288}{25} \\ -\frac{288}{25} & -\frac{216}{25} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{384}{25} v_1 - \frac{288}{25} v_2 \\ -\frac{288}{25} v_1 - \frac{216}{25} v_2 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{matrix} \mathrm{I}) \\ \mathrm{II}) \end{matrix}$$

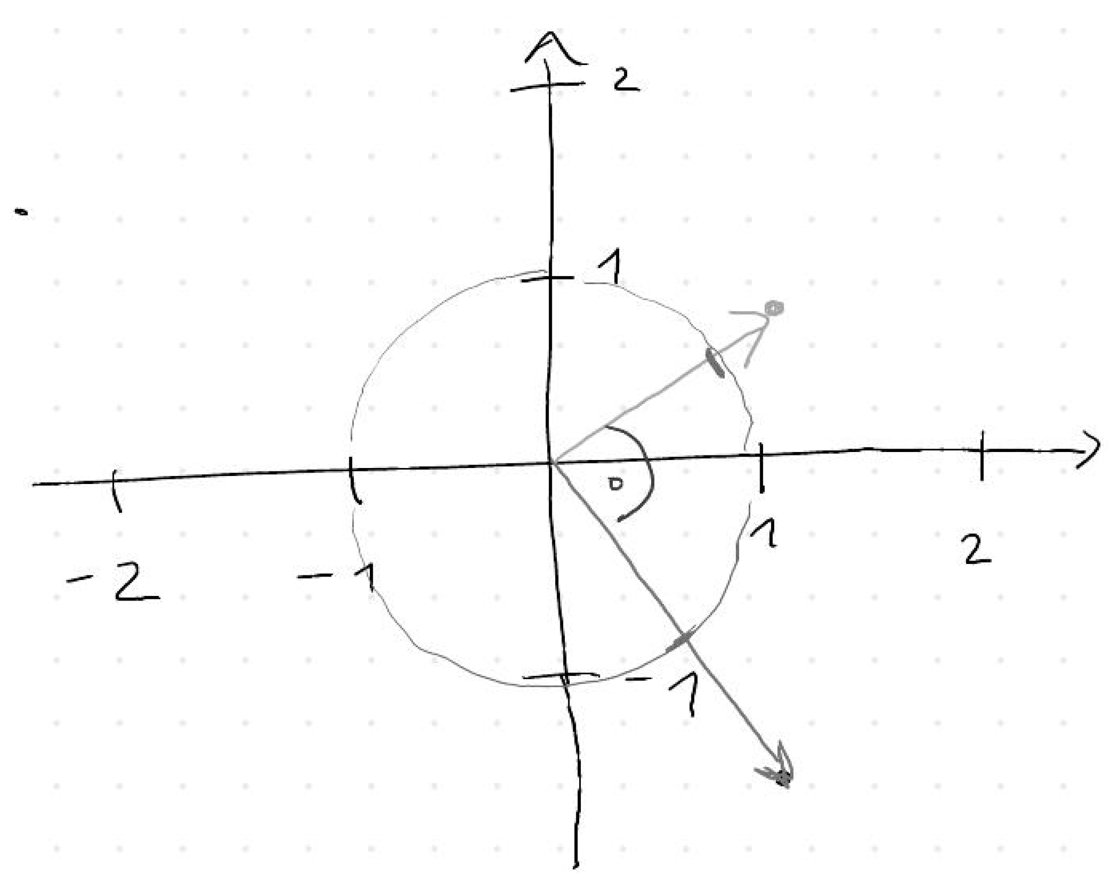
$\mathrm{I}) \Rightarrow \quad v_2 = -\frac{4}{3} v_1 \qquad \mathrm{II}) \Rightarrow \quad -\frac{288}{25} v_1 + \frac{216}{25} \cdot \frac{4}{3} v_1 = 0$

$\Rightarrow$ Eigenvektoren zu λ_1 haben die Form $\begin{pmatrix} v_1 \\ -\frac{4}{3} v_1 \end{pmatrix}$, $v_1 \in \mathbb{R}$.

Wir wählen z.B. $v_1 = 1$ & setzen $V_1 = \begin{pmatrix} 1 \\ -4/3 \end{pmatrix}$.

Für λ_2 gehen wir genauso vor. Die Eigenvektoren haben die Form $\begin{pmatrix} v_1 \\ 3/4\, v_1 \end{pmatrix}$. Wir wählen $V_2 = \begin{pmatrix} 1 \\ 3/4 \end{pmatrix}$. Wir wissen schon, dass V_1, V_2 orthogonal sind, können dies aber auch nachrechnen:

$$\langle V_1, V_2 \rangle = \left\langle \begin{pmatrix} 1 \\ -4/3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3/4 \end{pmatrix} \right\rangle = 1 \cdot 1 - \frac{4}{3} \cdot \frac{3}{4} = 1 - 1 = 0.$$



Wir normalisieren die Vektoren noch:

$$W_1 = \frac{V_1}{\|V_1\|} = \frac{\begin{pmatrix} 1 \\ -4/3 \end{pmatrix}}{\sqrt{1^2 + (-4/3)^2}} = \frac{3}{5} \cdot \begin{pmatrix} 1 \\ -4/3 \end{pmatrix} = \begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix}$$

$$W_2 = \frac{V_2}{\|V_2\|} = \frac{\begin{pmatrix} 1 \\ 3/4 \end{pmatrix}}{\sqrt{1^2 + (3/4)^2}} = \frac{4}{5} \begin{pmatrix} 1 \\ 3/4 \end{pmatrix} = \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}.$$

Dies ergibt die Matrix $V = \begin{pmatrix} | & | \\ W_1 & W_2 \\ | & | \end{pmatrix} = \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}$

2. Matrix U berechnen

a) Wir setzen zunächst $u_i = \dfrac{Aw_i}{\|Aw_i\|}$, also

$$Aw_1 = \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \\ 0 \end{pmatrix} \Rightarrow u_1 = \frac{Aw_1}{\|Aw_1\|} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$Aw_2 = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix} \Rightarrow u_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

b) Nun müssen wir U zu einer 3×3-Matrix ergänzen. Wir machen dies, indem wir $\{u_1, u_2\}$ zu einer ONB des $\mathbb{R}^3$ ergänzen.
(Man sieht hier eigentlich sofort, was zu tun ist, aber wir wenden nun das Gram-Schmidt-Verfahren an, um es einmal gesehen zu haben).

Wir starten, indem wir zu $\{u_1, u_2\}$ einen beliebigen linear unabhängigen Vektor hinzufügen, um eine Basis des $\mathbb{R}^3$ zu erhalten.
Wir setzen

$$u_3 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

Für $\{b_1, b_2, b_3\} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\}$

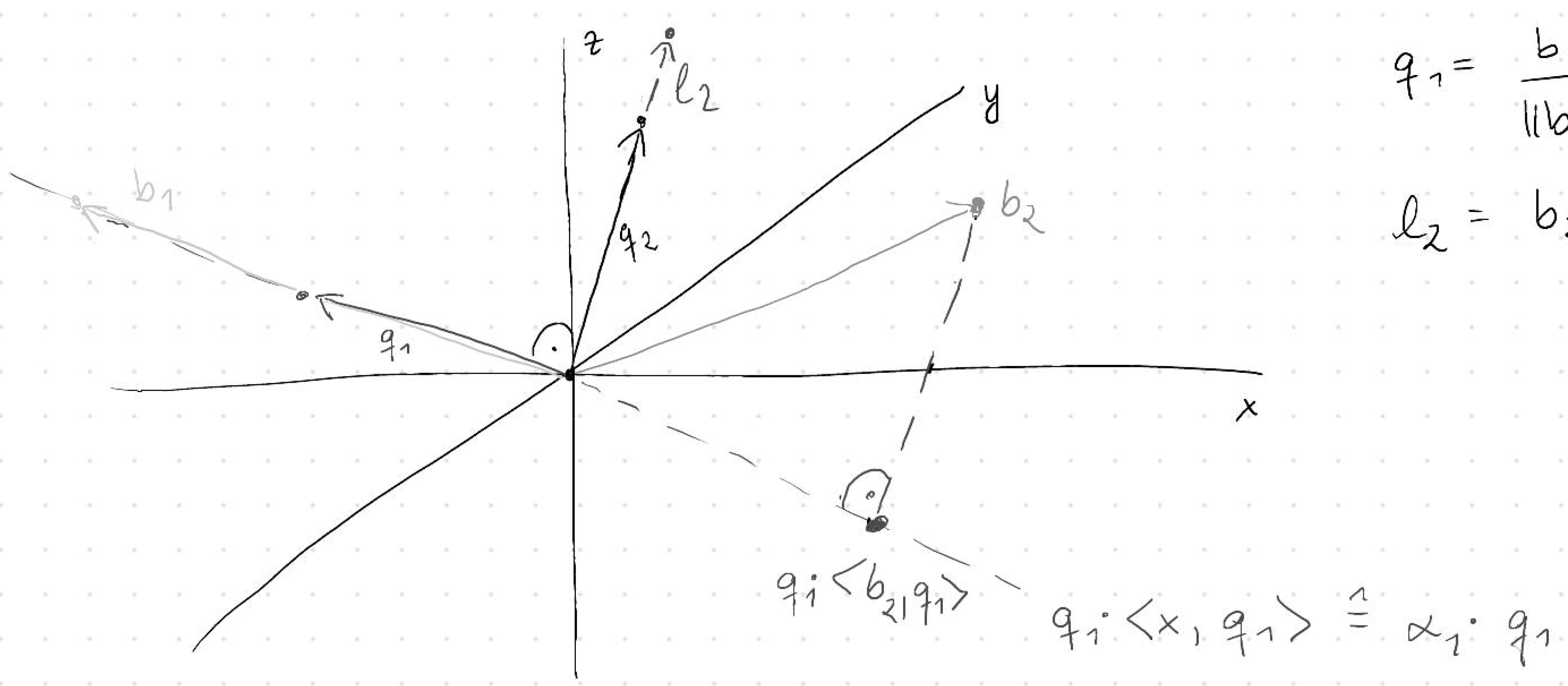
wenden wir Gram-Schmidt an.

1. ✓ $q_1 = b_1$
2. $$l_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} - \underbrace{\left\langle \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle}_{=0} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

Definition B.2.25. Sei $B = \{b_1, \ldots, b_n\}$ eine Basis im $\mathbb{K}^n$. Das *Gram-Schmidtsche Orthogonalisierungsverfahren* konstruiert eine ONB $\{q_1, \ldots, q_n\}$, indem die folgenden Schritte durchgeführt werden:

1. Normiere den ersten Basisvektors:
$$q_1 := \frac{b_1}{\|b_1\|}.$$
2. Fälle das Lot von b_2 auf die von q_1 erzeugte Gerade:
$$l_2 := b_2 - \langle b_2, q_1 \rangle q_1.$$
3. Normiere das Lot:
$$q_2 := \frac{l_2}{\|l_2\|}.$$
4. Fälle das Lot von b_k auf die von $q_1, \ldots, q_{k-1}$ erzeugte Ebene für $k = 3, \ldots, n$:
$$l_k := b_k - \langle b_k, q_1 \rangle q_1 - \ldots - \langle b_k, q_{k-1} \rangle q_{k-1}$$
5. Normiere das Lot l_k, $k = 3, \ldots, n$:
$$q_k := \frac{l_k}{\|l_k\|}.$$

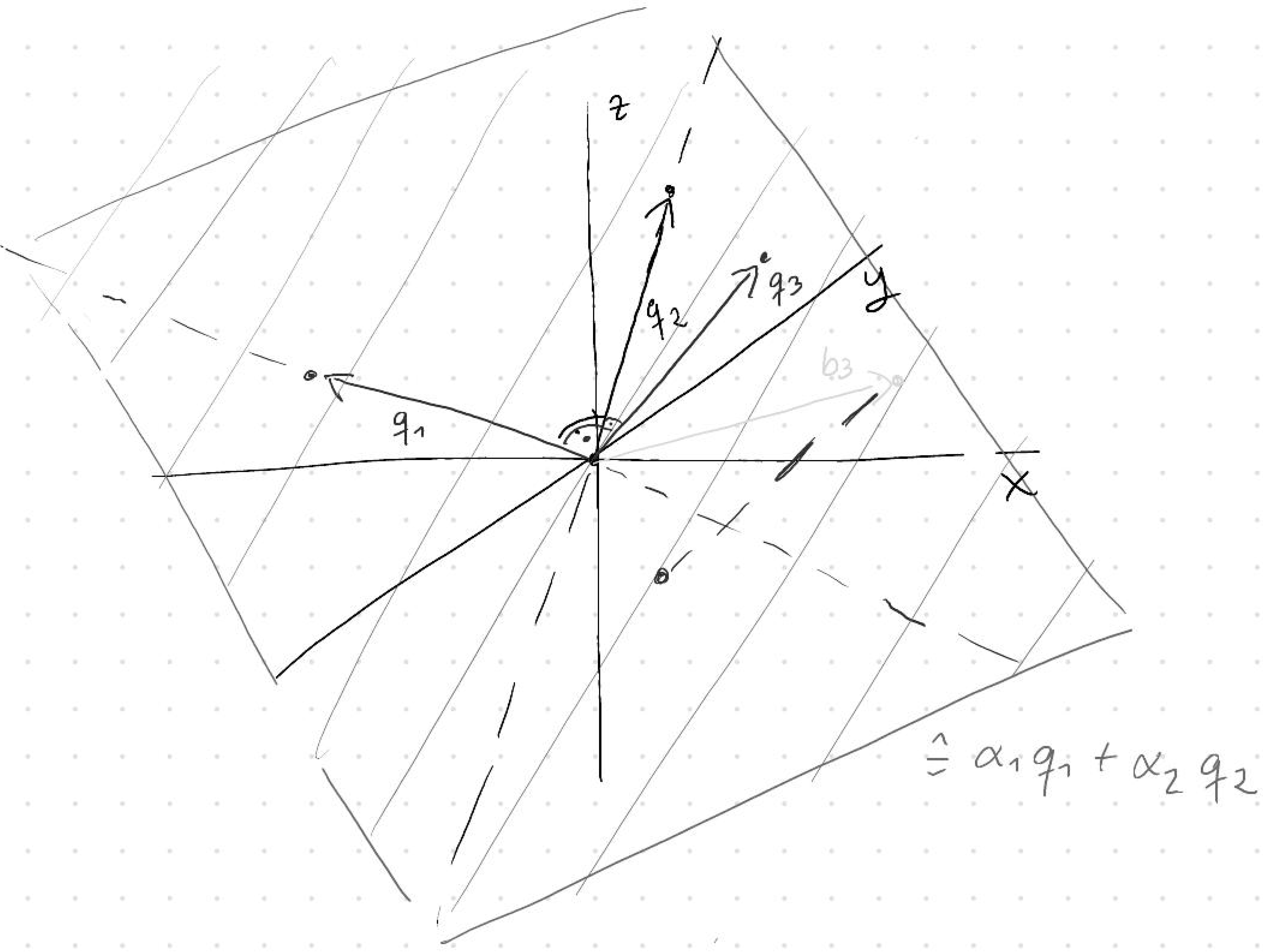
3. $q_2 = l_2 = b_2$ ✓



$$q_1 = \frac{b_1}{\|b_1\|}$$

$$l_2 = b_2 - q_1 \cdot \langle b_2, q_1 \rangle$$

4. $$l_3 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$
5. $$q_3 = \frac{l_3}{\|l_3\|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$



Wir setzen jetzt

$$U = \begin{pmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Σ berechnen

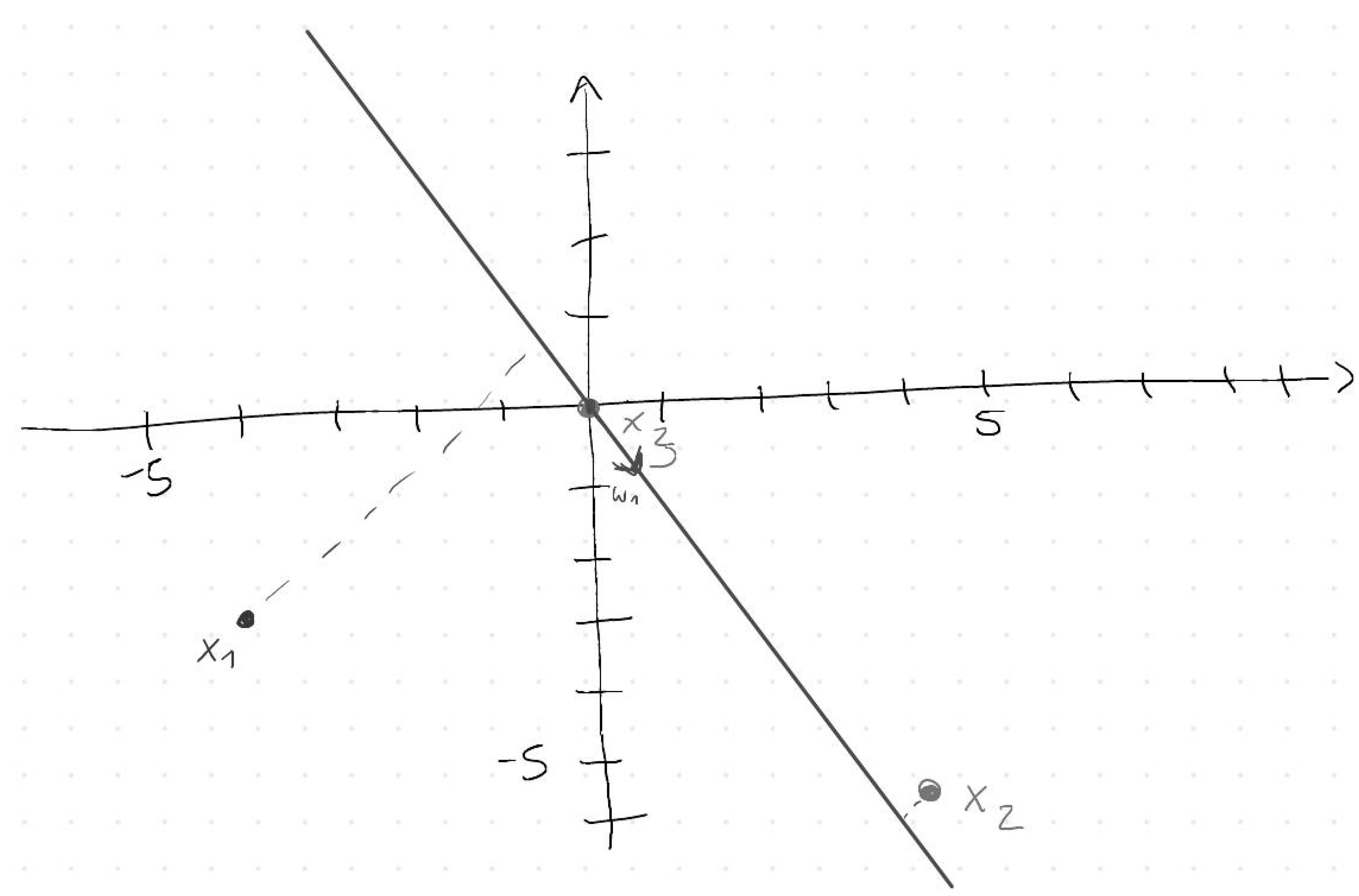
$$\sigma_{ij} = \begin{cases} \sqrt{\lambda_i} & i = j \\ 0 & \text{sonst} \end{cases} \Rightarrow \Sigma = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{49} & 0 \\ 0 & \sqrt{25} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix}$$

Insgesamt:

$$A = \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 7 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}^T}_{V^T}$$

⇝ Aufgabe 2: Wir fassen A als Sammlung von Datenpunkten im $\mathbb{R}^2$ auf (x^i = ite Zeile von A).

$$A = \begin{pmatrix} -4 & -3 \\ \frac{21}{5} & -\frac{28}{5} \\ 0 & 0 \end{pmatrix}$$



Die Singulärwertzerl. kann als Regression verstanden werden. Die Achsen (Spalten von V) gehen dabei immer durch den Ursprung.

Falls die Daten zentriert sind (d.h. der Mittelwert liegt in 0), dann entsprechen diese Achsen genau der Regressionsgeraden aus der linearen Regression: Wir haben

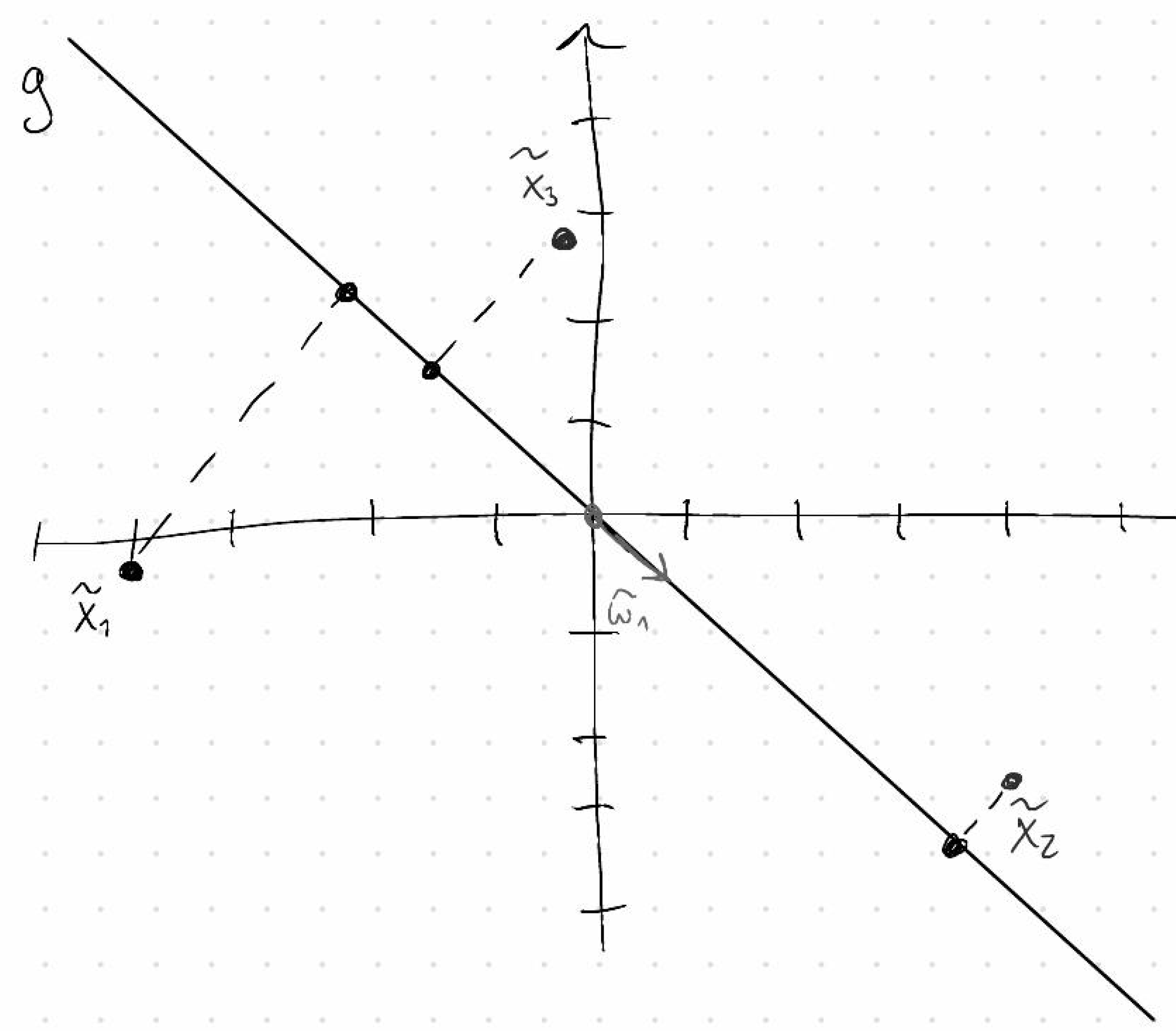
$$\tilde{A} = \begin{pmatrix} -4-\mu_x & -3-\mu_y \\ \frac{21}{5}-\mu_x & -\frac{28}{5}-\mu_y \\ 0-\mu_x & 0-\mu_y \end{pmatrix}$$

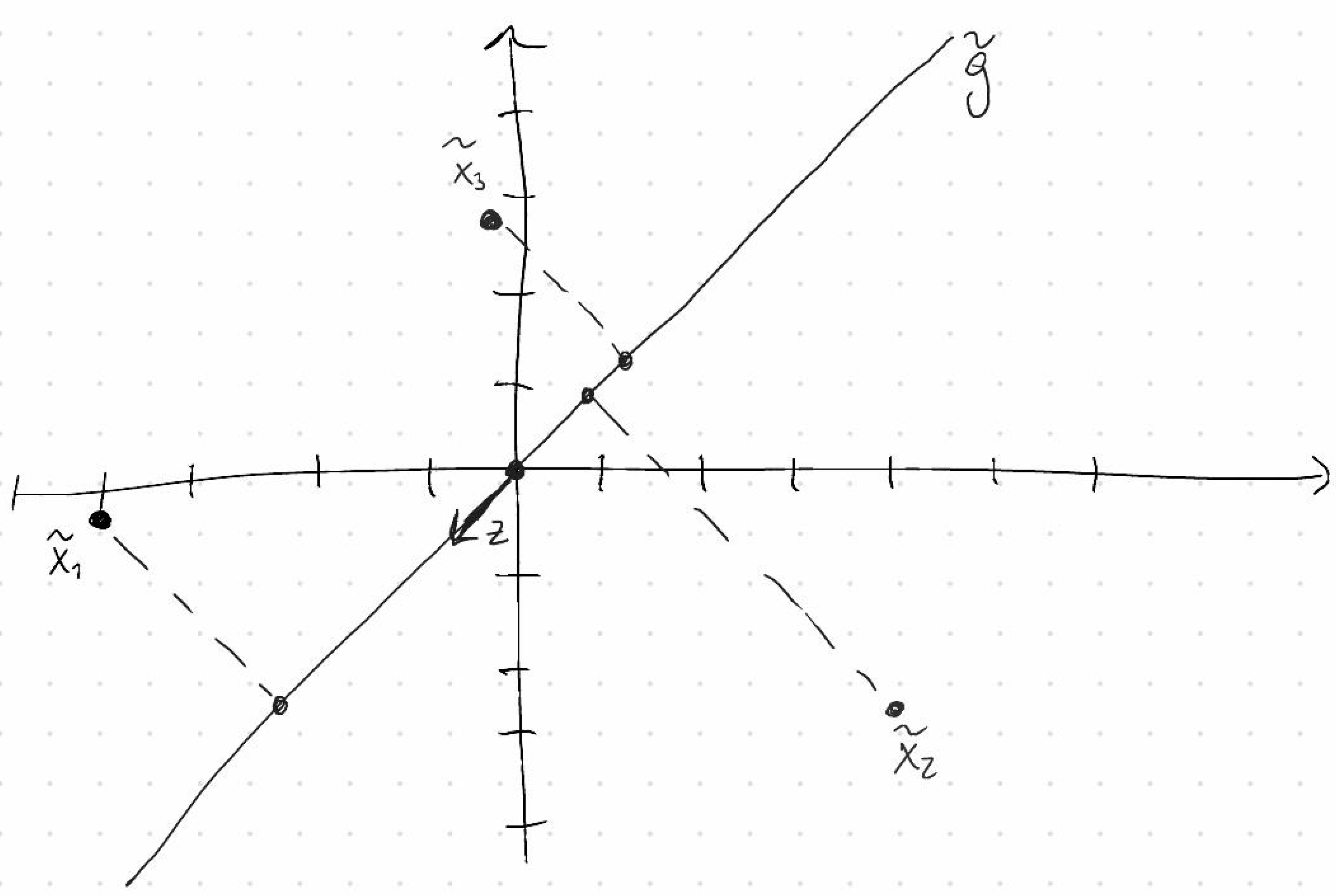
für $\mu_x = \frac{\left(-4+\frac{21}{5}+0\right)}{3} = 1/15$

$\mu_y = \frac{\left(-3-\frac{28}{5}+0\right)}{3} = -\frac{43}{15}$

$$= \begin{pmatrix} -\frac{61}{15} & -\frac{2}{15} \\ \frac{62}{15} & -\frac{41}{15} \\ -\frac{1}{15} & \frac{43}{15} \end{pmatrix}$$

$$\approx \begin{pmatrix} -4 & -0{,}1 \\ 4.1 & -2{,}7 \\ -0.1 & 2.9 \end{pmatrix}$$





Idee: Jede Gerade entspricht einen Vektor z (der Länge 1)
Die projizierten Punkte sind

$$\hat{x}_i = \langle \tilde{x}_i, z\rangle \cdot \frac{z}{\|z\|^2}$$

a) Der quadratische Abstand wird

$$\|x_i - \hat{x}_i\|^2 = \| x_i - \langle x_i, z\rangle \cdot \frac{z}{\|z\|^2}\|^2 \overset{\|z\|=1}{=} \| x_i - \langle x_i, z\rangle \cdot z\|^2$$

$$= \langle x_i - \langle x_i, z\rangle \cdot z,\ x_i - \langle x_i, z\rangle \cdot z\rangle$$

$$\overset{*}{=} \langle x_i, x_i\rangle - 2\langle x_i, \langle x_i, z\rangle z\rangle + \langle \langle x_i, z\rangle z, \langle x_i, z\rangle z\rangle$$

$$= \underbrace{\langle x_i, x_i\rangle}_{=\|x_i\|^2} - 2\underbrace{\langle x_i, z\rangle \langle x_i, z\rangle}_{=\langle x_i, z\rangle^2} + \langle x_i, z\rangle^2 \underbrace{\langle z, z\rangle}_{=1}$$

$$= \|x_i\|^2 - \langle x_i, z\rangle^2$$

$$\Rightarrow \min_z \sum_{i=1}^n \|x_i - \tilde{x}_i\|^2 = \min_z \sum_{i=1}^n \|x_i\|^2 - \langle x_i, z\rangle^2$$

$*$

$\langle a-b, a-b\rangle$
$= \langle a, a-b\rangle - \langle b, a-b\rangle$
$= \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle$
$= \langle a,a\rangle - 2\langle a,b\rangle + \langle b,b\rangle$

$\langle \gamma a, \lambda b\rangle = \gamma\langle a, \lambda b\rangle = \gamma\lambda\langle a,b\rangle$

$$= \min_z \underbrace{\sum_{i=1}^n \|x_i\|^2}_{\text{unabh. von } z} - \sum_{i=1}^n \langle x_i, z\rangle^2 = \sum_{i=1}^n \|x_i\|^2 + \min_z \left(-\sum_{i=1}^n \langle x_i, z\rangle^2\right)$$

$$= \sum_{i=1}^n \|x_i\|^2 - \left(\max_z \sum_{i=1}^n \langle x_i, z\rangle^2\right) \quad (*)$$

b) Def. Operatornorm! $\Rightarrow \max_z \sum_{i=1}^n \langle x_i, z\rangle^2 = \max_z \|\tilde{A} z\|^2 = \sigma^2$

c) „Die erste Spalte $\tilde{w}_1$ von V löst das Optimierungsproblem $(*)$"

Idee: Man zeigt, dass $\|\tilde{A} \cdot \tilde{w}_1\| = \sigma^2$

In unserem Beispiel ist $\tilde{w}_1 \approx \begin{pmatrix} 0.9 \\ -0.4 \end{pmatrix}$, was der Gerade g oben entspricht.