Lineare Algebra

9. Übungsstunde

Steven Battilana
stevenb👊student.ethz.ch
battilana.uk/teaching

November 21, 2019

1 Skalarprodukt

Definition. Eine Abbildung $B : \mathbb{E}^n \times \mathbb{E}^n \to \mathbb{E}$, $(v,w) \mapsto B(v,w)$ heisst ***Bilinearfrom*** (BLF) auf $\mathbb{E}^n$, falls die Linearität in beiden Argumenten erfüllt ist: $\forall u,v,w \in \mathbb{E}^n$, $\forall \lambda \in \mathbb{E}$:

- $B(u+v,w) = B(u,w) + B(v,w), \quad B(\lambda u, v) = \lambda B(u,v)$
- $B(u,v+w) = B(u,v) + B(u,w), \quad B(u,\lambda v) = \lambda B(u,v)$.

Definition.
B eine BLF die auf $\mathbb{R}$ abbildet, heisst *symmetrisch*, falls $B(v,w) = B(w,v)$ und *alternierend*, falls $B(v,w) = -B(w,v)$, für $\forall v,w \in \mathbb{E}^n$.

Definition. Eine Abbildung $B : \mathbb{E}^n \times \mathbb{E}^n \to \mathbb{E}$, $(v,w) \mapsto B(v,w)$ heisst *sesquilinear* auf $\mathbb{E}^n$, falls die Linearität im zweiten Argumente erfüllt ist und im ersten Argument semilinear ist: $\forall u,v,w \in \mathbb{E}^n$, $\forall \lambda \in \mathbb{E}$:

- $B(u+v,w) = B(u,w) + B(v,w), \quad B(\lambda u, v) = \overline{\boldsymbol{\lambda}} B(u,v)$
- $B(u,v+w) = B(u,v) + B(u,w), \quad B(u,\lambda v) = \lambda B(u,v)$.

Definition.
B eine BLF die auf $\mathbb{C}$ abbildet, heisst *hermitesch*, falls $B(v,w) = \overline{B(w,v)}$.

Definition.
Sei B eine BLF die auf $\mathbb{C}$ abbildet, heisst *positiv definit*, wenn

$$B(v,v) > 0 \quad \text{für jedes } v \in \mathbb{E}^n \text{ mit } \ v \neq 0.$$

Bemerkung.
Eine symmetrische, positiv definite BLF heisst euklidisches Skalarprodukt.

Definition.
Ein ***Skalarprodukt*** ist eine Abbildung $\langle \cdot, \cdot \rangle : \mathbb{E}^n \times \mathbb{E}^n \to \mathbb{E}, (v,w) \mapsto \langle v,w \rangle = \sum_{k=1}^n \overline{v_k} w_k$ so dass: $\forall u,v,w \in \mathbb{E}^n$, $\forall \lambda \in \mathbb{E}$:

(i) $\langle v + \lambda w, u \rangle = \langle v,u \rangle + \overline{\lambda} \langle w,u \rangle$ (Bilinearität bzw. Sesquilinearität)

(ii) $\langle v,w \rangle = \overline{\langle w,v \rangle}$ (symmetrie bzw. hermitesch)

(iii) $\langle v,v \rangle \geq 0$ und $\langle v,v \rangle = 0 \Leftrightarrow v = 0$ (positiv semidefinit)

Definition.
Die Länge oder *euklidische Norm* eines Vektors $x \in \mathbb{E}^n$ ist die nichtnegative reelle Zahl $\|x\|$ definiert durch

$$\|x\| := \sqrt{\langle x,x \rangle} = \sqrt{x^H x} = \sqrt{\sum_{k=1}^n |x_k|^2}.$$

Bemerkung.
Sei $x,y \in \mathbb{E}^n, \alpha \in \mathbb{E}$, dann gilt:

$$\langle x,y \rangle = \|x\| \, \|y\| \cos(\alpha).$$

Definition.
Sei $\mathcal{B} = \{b_i\}_{i \in I}$ eine Basis von V, $\langle \cdot, \cdot \rangle$ Skalarprodukt auf V.

- $\mathcal{B}$ heisst *orthogonal* $\quad \Leftrightarrow \quad \langle b_k, b_l \rangle = 0, \forall k \neq l$
  In Worten: Basisvektoren "stehen $\perp$ aufeinander".
- $\mathcal{B}$ heisst *orthonormal* (ONB) $\quad \Leftrightarrow \quad \langle b_k, b_l \rangle = \delta_{kl} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$
  In Worten: Basisvektoren "stehen $\perp$ aufeinander" *und* haben Länge 1.

Bemerkung.
Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, $\mathcal{B} = \{b_i\}_{i \in I}$ Basis von V. Per Definition von Basis kann man jedes $v \in V$ schreiben als $v = \sum_{i \in I} \alpha_i b_i$. Falls $\mathcal{B}$ eine ONB ist, so ist $\alpha_i = \langle b_i, v \rangle, \ \forall i \in I$,
$\Rightarrow$ Koordinaten von v bezüglich $\mathcal{B}$ sind $\langle b_i, v \rangle$.

2 Das äussere Produkt und die orthogonale Projection

Bemerkung.

- innere Produkt: $\forall x, y \in \mathbb{E}^n$: $\qquad \langle x, y \rangle = x^H y \in \mathbb{E}$
- äussere Produkt: $\forall x \in \mathbb{E}^m, \forall y \in \mathbb{E}^n$: $\qquad xy^H \in \mathbb{E}^{m \times n}$

Satz.
Die *orthogonale Projektion* $P_Y x$ von $x \in \mathbb{E}^n$ auf die durch y und Ursprung induzierte Gerade ist:

$$P_Y x := \frac{1}{\|y\|^2} y y^H x = u u^H x \text{ mit } u := \frac{y}{\|y\|}$$

(i) y-Richtung: $\quad P_Y x = \alpha y, \ \alpha \in \mathbb{E}$
($P_Y x$ ist ein Punkt auf y).

(ii) x-Richtung: $\quad (x - P_Y x) \perp y$.

Bemerkung. *Projektionsmatrix*

- $P_Y x := \frac{1}{\|y\|^2} y y^H x = u u^H x$ mit $u := \frac{y}{\|y\|}$
- $P_Y^H = P_Y$ (symmetrisch/hermitesch)
- $P_Y^2 = P_Y$ (idempotent; dies ich eine alternative Definition von einer Projektion)

3 Gram-Schmidt Verfahren

Satz.
Zu jedem Vektorraum endlicher oder zumindest abzählbarer Dimension existiert einen ONB. (d.h. ist konstruierbar mit Gram-Schmidt)

Gram-Schmidt Verfahren
Sei V ein euklidischer/unitärer Vektorraum mit $\dim(V) = n < \infty$
Gegeben: $W = \text{span}\{v_1, ..., v_m\} \subset V$, $\{v_i\}_{i=1}^{m}$ Basis von W ist linear unabhängig
Gesucht 1: $w_1, ..., w_m$ so dass $W = \text{span}\{w_1, ..., w_m\}$ und $\{w_i\}_{i=1}^{\boldsymbol{m}}$ ONB von W ist orthonormal
Gesucht 2: $w_1, ..., w_m, w_{m+1}, ..., w_n$ so dass $V = \text{span}\{w_1, ..., w_n\}$ und $\{w_i\}_{i=1}^{\boldsymbol{n}}$ ONB von V ist orthonormal.

1. $w_1 = \frac{v_1}{\|v_1\|}$ mit $\|v_1\| = \sqrt{\langle v_1, v_1\rangle} \quad \Rightarrow W^{(1)} = \{w_1\}$
2. für $k \in \{2, ..., m\}$
   - Wähle $v \in V \setminus \text{span}\{W^{(k-1)}\}$
   - $w'_k = v - \sum\limits_{i=1}^{k-1} \langle v, w_i\rangle w_i = v - \langle v, w_1\rangle w_1 - ... - \langle v, w_{k-1}\rangle w_{k-1}$
   - $w_k = \frac{w'_k}{\|w'_k\|}$ mit $\|w'_k\| = \sqrt{\langle w'_k, w'_k\rangle}$
   - $W^{(k)} = W^{(k-1)} \cup \{w_k\} = \{w_1, ..., w_k\}$
3. $W^{(m)}$ ist eine ONB für W.
4. Falls man $W^{(m)}$ zu eine ONB von V erweitern will:
   Für $k \in \{m+1, ..., n\}$
   - Wähle $v \in V \setminus \text{span}\{W^{(k-1)}\}$
   - $w'_k = v - \sum\limits_{i=1}^{k-1} \langle v, w_i\rangle w_i$
   - $w_k = \frac{w'_k}{\|w'_k\|}$ mit $\|w'_k\| = \sqrt{\langle w'_k, w'_k\rangle}$
   - $W^{(k)} = W^{(k-1)} \cup \{w_k\}$
5. $W^{(n)}$ ist eine ONB für V.

Beispiel 1:

Gegeben: Vektorraum $\mathcal{P}_2$ über $\mathbb{R}$ mit Skalarprodukt $\langle p,q\rangle_{\mathcal{P}} = \int\limits_{-1}^{1} p(t)q(t)dt$ und Monombasis $\mathcal{B} = \{1, t, t^2\}$. Hier: $V = W = \mathcal{P}_2$
Gesucht: Eine ONB die bezüglich $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ orthogonal ist $\quad\Rightarrow\quad$ wir müssen $1, t, t^2$ bezüglich $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ orthonormieren.

Lösung:

1.

$$w_1 = \frac{1}{\|1\|_{\mathcal{P}}} = \frac{1}{\sqrt{2}}, \quad \text{weil} \quad \|1\|_{\mathcal{P}} = \sqrt{\int_{-1}^{1} 1dt} = \sqrt{2}$$

$$\Rightarrow W^{(1)} = \left\{\frac{1}{\sqrt{2}}\right\}$$

2. $k = 2:$

$$\begin{aligned} w_2' &= t - \langle t, \frac{1}{\sqrt{2}}\rangle \frac{1}{\sqrt{2}} \\ &= t - \int_{-1}^{1} t \cdot \frac{1}{\sqrt{2}} dt \\ &= t - \left[\frac{1}{2\sqrt{2}} t^2\right]_{-1}^{1} \\ &= t - 0 \\ &= t \end{aligned}$$

$$w_2 = \frac{t}{\|t\|_{\mathcal{P}}} = \sqrt{\frac{3}{2}}t \quad \text{weil} \quad \|t\|_{\mathcal{P}} = \sqrt{\int_{-1}^{1} t^2 dt} = \sqrt{\left[\frac{1}{3}t^3\right]_{-1}^{1}} = \sqrt{\frac{2}{3}}$$

$$W^{(2)} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}t\right\}$$

3. $k = 3:$

$$\begin{aligned} w_3' &= t^2 - \langle t^2, \frac{1}{\sqrt{2}}\rangle \frac{1}{\sqrt{2}} - \langle t^2, \sqrt{\frac{3}{2}}t\rangle \sqrt{\frac{3}{2}}t \\ &= t^2 - \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{1}{\sqrt{2}} \cdot t^2 dt - \sqrt{\frac{3}{2}}t \int_{-1}^{1} \sqrt{\frac{3}{2}}t \cdot t^2 dt \\ &= t^2 - \frac{1}{\sqrt{2}} \left[\frac{1}{3\sqrt{2}} t^3\right]_{-1}^{1} - \sqrt{\frac{3}{2}}t \left[\sqrt{\frac{3}{2}}\frac{1}{4}t^4\right]_{-1}^{1} \\ &= t^2 - \frac{1}{3} - 0 \\ &= t^2 - \frac{1}{3} \end{aligned}$$

$$w_3 = \frac{w_3'}{\|w_3'\|_{\mathcal{P}}} = \sqrt{\frac{45}{8}}(t^2 - \frac{1}{3}) \quad \text{weil} \quad \|w_3'\|_{\mathcal{P}} = \int_{-1}^{1} (t^2 - \frac{1}{3})^2 dt = \left[\frac{t^5}{5} - \frac{2t^2}{9} + \frac{t}{9}\right]_{-1}^{1} = \frac{8}{45}$$

$$W^{(3)} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}t, \sqrt{\frac{45}{8}}(t^2 - \frac{1}{3})\right\} \qquad \text{ist ONB für } \mathcal{P}_2.$$

4 QR-Zerlegung

Bemerkung.
Sei $A \in \mathbb{E}^{m\times n}$. Angenommen $\text{rang}(A) = n \leq m$ (voller Rang). Dann existiert eine orthogonale/unitäre (Matrix $Q = (Q_1|Q_2)$ mit $Q \in \mathbb{E}^{m\times m},\ Q_1 \in \mathbb{E}^{m\times n},\ Q_2 \in \mathbb{E}^{m\times(m-n)}$ und es existiert eine rechte Dreiecksmatrix $R_1 \in \mathbb{E}^{n\times n}$ und mit $\mathbf{0} \in \mathbb{E}^{(m-n)\times n}$, so dass

$$A = Q \cdot \begin{pmatrix} R_1 \\ \hline \mathbf{0} \end{pmatrix} =: QR$$

Bemerkung
Da $\begin{pmatrix} R_1 \\ \hline \mathbf{0} \end{pmatrix}$, ist $A = QR = Q_1R_1$.

Bemerkung.(Zusammenhang QR-Zerlegung und kleinste Quadrate)
Für das Residuum r gilt, dann $r = y - \sum_{i=1}^{n} \langle q_i, y\rangle q_i$, wobei $\{q_i\}_{i=1}^n$ sie Spalten von Q bezeichnet.

Bemerkung.
Falls das betrachtete Skalarprodukt das Standard-Skalarprodukt ist, gilt:

$$\begin{aligned} Q_1^H Q_1 &= \mathbb{1} \\ Ax = y &\overset{A=Q_1R_1}{\iff} Q_1R_1x = y \\ &\Leftrightarrow R_1x = Q_1^H y \\ &\Leftrightarrow x = R_1^{-1}Q_1^H y \end{aligned}$$

Für ein anderes Skalarprodukt erggibt sich eine andere Fromel für x!

QR-Zerlegung
Gegeben: $A \in \mathbb{E}^{m\times n},\ m > n,\ \langle\cdot,\cdot\rangle$ ein Skalarprodukt
Gesucht: Q_1, R_1, so dass $A = Q_1R_1$.

Q1: Seien $a_1, ..., a_n$ Spalten von A. Wende Gram-Schmidt Verfahren bezüglich $\langle\cdot,\cdot\rangle$ auch $\{a_i\}_{i=1}^n$ an. Das liefert dann $\{q_1, ..., q_n\}$.
$\Rightarrow (q_1\ |...|q_n) = Q_1 \in \mathbb{E}^{m\times n}$

R1:
1.
$$r_{ij} = \begin{cases} 0, & i > j \\ \langle q_i, a_j\rangle, & i < j \\ \|\widetilde{q_i}\| = \sqrt{\langle \widetilde{q_i}, \widetilde{q_i}\rangle}, & i = j \end{cases}$$
wobei $\widetilde{q_i}$ den i-ten orthogonalisierten, aber noch nicht normierten Vektor bezeichnet.

2. $R_1 = (r_{ij})_{1\leq i,j\leq n}$

Beispiel 2:
Gegeben:
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 3 \\ 2 & -2 \end{pmatrix}, \quad \Rightarrow \quad m = 3 > 2 = n, \quad y = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}$$

Gesucht: Bestimme mit Hilfe der QR-Zerlegung $x = (x_1, x_2)^T$, so dass $\|r\|_2^2 = \|y - Ax\|_2^2$ minimal ist bezüglich dem Standard-Skalarprodukt $\langle x, y\rangle = x^H y$.

Lösung:

(Q1)

$$
\begin{aligned}
q_1 &= \frac{a_1}{\|a_1\|} = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \\
\widetilde{q_2} &= a_2 - \langle q_1, a_2 \rangle q_1 \\
&= \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} - \frac{1}{\sqrt{5}} \frac{1}{\sqrt{5}} \langle \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \rangle \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} - \frac{1}{5} \cdot (-1) \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ \frac{16}{5} \\ -\frac{8}{5} \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 5 \\ 16 \\ -8 \end{pmatrix} \\
q_2 &= \frac{\widetilde{q_2}}{\|\widetilde{q_2}\|} = \frac{1}{5} \frac{5}{\sqrt{345}} \begin{pmatrix} 5 \\ 16 \\ -8 \end{pmatrix} = \frac{1}{\sqrt{345}} \begin{pmatrix} 5 \\ 16 \\ -8 \end{pmatrix} \\
\Rightarrow \quad Q_1 &= \begin{pmatrix} 0 & \frac{5}{\sqrt{345}} \\ \frac{1}{\sqrt{5}} & \frac{16}{\sqrt{345}} \\ \frac{2}{\sqrt{5}} & \frac{-8}{\sqrt{345}} \end{pmatrix}
\end{aligned}
$$

(R1)

$$
\begin{aligned}
r_{11} &= \|a_1\| = \sqrt{5} \\
r_{22} &= \|\widetilde{q_2}\| = \left(\frac{1}{5} \frac{1}{5} \langle \begin{pmatrix} 5 \\ 16 \\ -8 \end{pmatrix}, \begin{pmatrix} 5 \\ 16 \\ -8 \end{pmatrix} \rangle \right)^{\frac{1}{2}} = \left(\frac{1}{25} (25 + 256 + 64) \right)^{\frac{1}{2}} = \frac{\sqrt{345}}{5} \\
r_{12} &= \langle q_1, a2 \rangle = \langle \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \rangle = -\frac{1}{\sqrt{5}}
\end{aligned}
$$

$$r_{21} = 0$$

$$
\begin{aligned}
\Rightarrow \quad R_1 &= \begin{pmatrix} \sqrt{5} & -\frac{1}{\sqrt{5}} \\ 0 & \frac{\sqrt{345}}{5} \end{pmatrix} \\
\Rightarrow \quad A &= Q_1 R_1
\end{aligned}
$$

$$
\begin{aligned}
Ax = y &\Rightarrow Q_1R_1x = y\\
&\Leftrightarrow R_1x = Q_1^Ty\\
&\Leftrightarrow x = R_1^{-1}Q_1^Ty\\
R_1^{-1} &= \frac{\sqrt{5}}{\sqrt{345}}\begin{pmatrix} \frac{\sqrt{345}}{5} & \frac{1}{\sqrt{5}}\\ 0 & \sqrt{5}\end{pmatrix}\\
Q_1^T &= \begin{pmatrix} 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{5}{\sqrt{345}} & \frac{16}{\sqrt{345}} & \frac{-8}{\sqrt{345}}\end{pmatrix}\\
x &= \begin{pmatrix} \frac{\sqrt{345}}{5} & \frac{1}{\sqrt{5}}\\ 0 & \sqrt{5}\end{pmatrix}\begin{pmatrix} 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{5}{\sqrt{345}} & \frac{16}{\sqrt{345}} & \frac{-8}{\sqrt{345}}\end{pmatrix}\begin{pmatrix}2\\2\\4\end{pmatrix}\\
&= \begin{pmatrix} \frac{5}{345} & \frac{85}{345} & \frac{130}{345}\\ \frac{25}{345} & \frac{80}{345} & -\frac{40}{345}\end{pmatrix}\begin{pmatrix}2\\2\\4\end{pmatrix}\\
&= \frac{1}{345}\begin{pmatrix} 5 & 85 & 130\\ 25 & 80 & -40\end{pmatrix}\begin{pmatrix}2\\2\\4\end{pmatrix}\\
&= \frac{1}{69}\begin{pmatrix} 1 & 17 & 26\\ 5 & 16 & -8\end{pmatrix}\begin{pmatrix}2\\2\\4\end{pmatrix}\\
&= \frac{1}{69}\begin{pmatrix} 2+34+104\\ 10+32-32\end{pmatrix}\\
&= \frac{1}{69}\begin{pmatrix} 140\\ 10\end{pmatrix}
\end{aligned}
$$

Korollar 10.10 (Eigenwerte werden später definiert)
Die 2-Norm-*Konditonszahl* (kurz: Kondition) einer Matrix $A \in \mathbb{E}^{n\times n}$ ist gegeben durch

$$\kappa_2 = \frac{\max\{|\omega|\}}{\min\{|\omega|\}},$$

wobei ω der Eigenwert von A ist.
Insbesondere ist die Konditionszahl immer grösser oder gleich 1.

Bemerkung.
Wir möchten immer eine möglichst kleine Kondition haben, da dann die Implementierung numerisch am stabilsten ist, d.h. grosse Kondition ist schlecht und unerwünscht.

Bemerkung.
Die QR-Zerlegung ist für $m \geq n$ und $\text{Rang}(A) = n$ eindeutig, wenn man die Vorzeichen der Diagonalelemente von der Dreiecksmatrix vorgibt.

Normalengleichung vs. QR-Zerlegung
Normalengleichung

Pro:
- "schön" um von Hand zu rechnen
- A^HA ist hermitesch positiv definit $\Rightarrow$ kann für Cholesky-Zerlegung ausgenutzt werden

Con:

- Wenn A schlecht Konditioniert ist, ist $A^H A$ quadratisch schlecht konditioniert. $\Rightarrow$ Cholesky liefert unbrauchbare Resultate (Rundungsfehler massiv verstärkt, Implementiereung instabil)

QR-Zerlegung

Pro:

- kann numerisch stabil implementiert werden (Verfahren hängt weniger von der Kondition ab)
- Q ist Projektion der Spalten von A

Con:

- "hässlich" zum von Hand rechnen (Wurzeln)