Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Gram-Schmidt-Verfahren, Orthogonale Projektionen

Die Themen heute sind

- Gram-Schmidt
- Orthogonale Projektionen
- Hilbert- und Banachräume
- Die unitären Gruppen $\mathrm{U}(n)$ und $\mathrm{SU}(n)$
- Hauptachsentransformation

Heute werden wir das Gram-Schmidt-Verfahren besprechen und dessen Anwendung auf orthogonale Projektionen behandeln. Anschließend befassen wir uns mit der Hauptachsentransformation.

Gram-Schmidt-Verfahren

Algorithmus. Sei V ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum mit einem Skalarprodukt $\langle\ ,\ \rangle$.

Input. Eine Familie $\{v_1, \ldots, v_n\}$ von linear unabhängigen Vektoren aus V.

Output. Eine orthonormale Familie $\{w_1, \ldots, w_n\}$ von Vektoren aus V, so dass

$$\text{Spann}(v_1, \ldots, v_k) = \text{Spann}(w_1, \ldots, w_k)\ \forall k \in \{1, \ldots n\} \text{ gilt.}$$

1. $w_1 = \frac{v_1}{\|v_1\|}$.
2. **for** k **from** 2 **to** n **do**
   a. $$u_k = v_k - \sum_{j=1}^{k-1} \langle w_j, v_k \rangle w_j$$
   b. $w_k = \frac{u_k}{\|u_k\|}$
3. **return** $\{w_1, \ldots w_n\}$

Beweis. Korrektheit: Wegen

$$\begin{aligned}\langle w_\ell, u_k\rangle &= \langle w_\ell, v_k\rangle - \sum_{j=1}^{k-1}\langle w_j, v_k\rangle\langle w_\ell, w_j\rangle \\ &= \langle w_\ell, v_k\rangle - \sum_{j=1}^{k-1}\langle w_j, v_k\rangle\delta_{\ell j} = \langle w_\ell, v_k\rangle - \langle w_\ell, v_k\rangle = 0\end{aligned}$$

steht u_k senkrecht auf $w_1, \dots, w_{k-1}$. Es gilt $u_k \neq 0$, da nach Vorraussetzung

$$v_k \notin \mathrm{Spann}(w_1, \dots, w_{k-1}) = \mathrm{Spann}(v_1, \dots, v_{k-1})$$

gilt. Der normierte Vektor w_k steht dann ebenfalls senkrecht auf $w_1, \dots, w_{k-1}$, und

$$\mathrm{Spann}(w_1, \dots, w_k) = \mathrm{Spann}(v_1, \dots v_{k-1}, u_k) = \mathrm{Spann}(v_1, \dots, v_k)$$

ergibt sich mit Induktion nach k. □

Orthogonale Projektionen

Satz. *Sei V ein Vektorraum mit einem Skalarprodukt $\langle\ ,\ \rangle$, $\{w_1, \ldots, w_n\}$ ein Orthonormalsystem aus Vektoren in V, und sei $v \in V$ ein weiterer Vektor. Dann ist*

$$w = \sum_{j=1}^{n} \langle w_j, v \rangle w_j$$

der Vektor in $W = \operatorname{Spann}(w_1, \ldots, w_n)$ mit dem kleinsten Abstand $\|v - w\|$ zu v.

Beweis. Es gilt $v - w \perp \operatorname{Spann}(w_1, \ldots, w_n)$, da

$$\begin{aligned}
\langle w_\ell, v - w \rangle &= \langle w_\ell, v \rangle - \sum_{j=1}^{k-1} \langle w_j, v \rangle \langle w_\ell, w_j \rangle \\
&= \langle w_\ell, v \rangle - \sum_{j=1}^{k-1} \langle w_j, v \rangle \delta_{\ell j} = \langle w_\ell, v \rangle - \langle w_\ell, v \rangle = 0.
\end{aligned}$$

Ist nun $\tilde{w} \in \mathrm{Spann}(w_1, \ldots, w_n)$ ein weiterer Punkt, dann gilt

$$\|v - \tilde{w}\|^2 = \|v - w\|^2 + \|w - \tilde{w}\|^2 \geqslant \|v - w\|^2$$

nach Pythagoras, da $v - w \perp w - \tilde{w}$, und Gleichheit gilt nur für $\tilde{w} = w$. $\square$

Wir können also auf endlich-dimensionale Unterräume eines Vektorraums mit Skalarprodukt orthogonal projizieren.

Beispiel. Im letzten Semester hatten wir den Vektorraum

$$V = \{f : \mathbb{R} \to \mathbb{C} \mid f \text{ ist stetig und } 2\pi - \text{periodisch}\}$$

betrachtet und für $f \in V$ mit den Fourier-Koeffizienten

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt$$

die Fourier-Reihe

$$\sum_{k=-\infty}^{\infty} c_k e^{ikt}$$

gebildet.

Fourier-Polynome

Jetzt sehen wir, dass die Funktionen

$$e_k : \mathbb{R} \to \mathbb{C}, \quad e_k(t) = e^{ikt}$$

ein Orthonormalsystem $\{e_k \mid k \in \mathbb{Z}\}$ in V bzgl. des Skalarprodukts

$$\langle g, f \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{g(t)} f(t) dt$$

bilden. Da $\overline{e^{ikt}} = e^{-ikt}$ gilt, ist das n-te Fourier-Polynom

$$\sum_{k=-n}^{n} c_k e^{ikt}$$

die orthogonale Projektion von f auf den Unterraum Spann$(e_{-n}, \ldots, e_n)$ der Fourier-Polynome vom Grad $\leqslant n$.

Ende Teil 1

Orthogonale Komplemente

Definition. Sei V ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum mit einem Skalarprodukt $\langle\ ,\ \rangle$ und $W \subset V$ ein Untervektorraum. Dann heißt

$$W^\perp = \{u \in V \mid \langle u, w\rangle = 0\ \forall w \in W\} = \{w \in V \mid w \perp W\}$$

das **orthogonale Komplement** von U in V.

Bemerkung. Hat W endliche Dimension und ist $\mathcal{B} = \{w_1, \ldots, w_k\}$ eine Orthonormalbasis von W, dann heißt die Abbildung

$$V \to W^\perp, \quad v \mapsto v - \sum_{j=1}^{k} \langle w_j, v\rangle w_j$$

die **orthogonale Projektion** auf $W^\perp$.

Dass der Vektor

$$u = v - \sum_{j=1}^{k} \langle w_j, v \rangle w_j$$

in $W^\perp \subset V$ liegt, folgt mit unserer Rechnung für die Projektion auf W. $u \in W^\perp$ mit kleinstem Abstand zu v:

$$\|v - u\| = \inf_{\tilde{u} \in W^\perp} \{\|v - \tilde{u}\|\} = \inf\{\|v - \tilde{u}\| \mid \tilde{u} \in W^\perp\}.$$

Satz. *Sei V ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum mit einem Skalarprodukt $\langle\, , \,\rangle$ endlicher Dimension und sei $W \subset V$ ein Untervektorraum. Dann gilt*

$$W \oplus W^\perp = V \text{ und } \dim W + \dim W^\perp = \dim V.$$

Beweis. Wir erhalten eine Orthonormalbasis von V indem wir Orthonormalbasen von W und $W^\perp$ zusammenfügen. □

Hilbert- und Banachräume

Ist $\dim V = \infty$, dann ist die Situation komplizierter, weil Ausdrücke der Art

$$\sum_{j=1}^{\infty} \langle w_j, v \rangle w_j$$

zunächst einmal keinen Sinn ergeben. Man benötigt dazu auch Analysis.

Definition. Sei (v_n) eine Folge von Vektoren in einem normierten Vektorraum V. Wir sagen, die Folge (v_n) konvergiert gegen einen Vektor $v \in V$, falls

$$\lim_{n\to\infty} \|v_n - v\| = 0$$

gilt. Ein Vektorraum V mit einem Skalarprodukt $\langle\ ,\ \rangle$ nennt man auch einen **Prä-Hilbertraum**. V ist ein **Hilbertraum**, wenn jede Cauchy-Folge (v_n) von Vektoren in V konvergiert. Analog ist ein **Banachraum** ein normierter Raum in der jede Cauchy-Folge konvergiert.

$\ell^2(\mathbb{R})$ und $\ell^2(\mathbb{C})$

Auf dem $\mathbb{R}$-Vektorraum der quadratsummierbaren Folgen

$$\ell^2(\mathbb{R}) = \{(x_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$$

ist durch

$$\langle (x_n), (y_n) \rangle = \sum_{n=1}^{\infty} x_n y_n \in \mathbb{R}$$

ein Skalarprodukt erklärt mit Norm

$$\|(x_n\| = (\sum_{n=1}^{\infty} |x_n|^2)^{\frac{1}{2}}.$$

Dies macht Sinn, da die Reihe für das Skalarprodukt (absolut) konvergiert:

$$\sum_{n=1}^{N} |x_n y_n| \leqslant (\sum_{n=1}^{N} |x_n|)^{\frac{1}{2}} (\sum_{n=1}^{N} |y_n|)^{\frac{1}{2}} \leqslant \|(x_n)\| \|(y_n)\| < \infty$$

gilt nach der Cauchy-Schwarzschen Ungleichung für $\mathbb{R}^N$.

$\ell^2(\mathbb{R})$ ist ein Hilbertraum, da für (v_k) eine Cauchy-Folge in $\ell^2(\mathbb{R})$, also $v_k = (v_{kn})_{n\in\mathbb{N}}$, die Folge der n-ten Glieder $(v_{kn})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$ bilden und $\mathbb{R}$ vollständig ist.

Beispiel. Ein Beispiel für einen Banachraum ist

$$V = C^0[a, b] = \{f : [a, b] \to \mathbb{R} \mid f \text{ ist stetig}\}$$

zusammen mit der Supremumsnorm

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [a, b]\}.$$

Es gilt $\|f\|_\infty < \infty$, da stetige Funktionen auf abgeschlossenen Intervallen ihr Maximum annehmen. Konvergenz bzgl. $\| \ \|_\infty$ ist gleichmäßige Konvergenz, und der Grenzwert f einer Funktionenfolge (f_n) ist bei gleichmäßiger Konvergenz erneut stetig.

Ende Teil 2

Isometrien

Definition. Seien V und W zwei $\mathbb{R}$- oder $\mathbb{C}$-Vektorräume mit Skalarprodukten. Ein Isomorphimus $f\colon V \to W$ ist eine **Isometrie**, wenn

$$\langle f(v_1), f(v_2)\rangle = \langle v_1, v_2\rangle \,\forall v_1, v_2 \in V$$

gilt.

Korollar. *Je zwei endlich-dimensionale $\mathbb{R}$- bzw. $\mathbb{C}$-Vektorräume gleicher Dimension sind isometrisch zueinander.*

Beweis. Nach dem Gram-Schmidt-Verfahren hat jeder endlich-dimensionale $\mathbb{C}$-Vektorraum V mit Skalarprodukt eine Orthonormalbasis. Ist $\mathcal{B} = \{v_1, \dots, v_n\}$ eine Orthonormalbasis eines $\mathbb{C}$ Vektorraums V, dann definiert

$$f\colon \mathbb{C}^n \to V,\ e_j \mapsto v_j$$

eine Isometrie zwischen $\mathbb{C}^n$ mit dem Standardskalarprodukt und V. Der Beweis für $\mathbb{R}$-Vektorräume ist analog. $\square$

Die unitären Gruppen $\mathrm{U}(n)$ und $\mathrm{SU}(n)$

Die Isometrie $f\colon \mathbb{C}^n \to V,\ e_j \mapsto v_j$ ist allerdings nicht eindeutig bestimmt. Eine Matrix $T \in \mathbb{C}^{n\times n}$ definiert eine Isometrie $\mathbb{C}^n \to \mathbb{C}^n, v \mapsto Tv$ bzgl. des Standardskalarprodukts genau dann, wenn die Spalten von T eine Orthonormalbasis von $\mathbb{C}^n$ bilden, also wenn

$$\overline{T}^t\, T = E$$

gilt. Solche Matrizen nennt man auch **unitär**.

Definition. Die Gruppe

$$\mathrm{U}(n) = \{\, T \in \mathrm{GL}(n,\mathbb{C}) \mid T^{-1} = \overline{T}^t\,\}$$

heißt **unitäre Gruppe** der Ordung n.

$$\mathrm{SU}(n) = \{\, T \in \mathrm{U}(n) \mid \det T = 1\}$$

heißt **spezielle unitäre Gruppe**.

Beispiel.

$$U(1) = \{(\lambda) \in GL(1,\mathbb{C}) \mid \overline{\lambda}\lambda = 1\} = \{z \in \mathbb{C} \mid |z| = 1\} \subset (\mathbb{C}^*, \cdot)$$

ist die Einheitskreislinie in $\mathbb{C}$.

$$\det\colon U(n) \to U(1),\ T \mapsto \det T$$

ist ein Gruppenhomomorphismus mit Kern $SU(n)$.
Bisher haben wir also folgende Matrixgruppen kennengelernt:

1. $SL(n,\mathbb{R}) \subset GL(n,\mathbb{R})$
2. $SL(n,\mathbb{C}) \subset GL(n,\mathbb{C})$
3. $SO(n) \subset O(n)$
4. $SU(n) \subset U(n)$

Ferner nennen wir Matrizen A mit $A^t = A$ symmetrisch und Matrizen B mit $B^t = \overline{B}$ hermitesch.

Satz. *Die Eigenwerte einer hermiteschen Matrix $B = (b_{ij}) \in \mathbb{C}^{n \times n}$ sind alle reell.*

Beweis. Sei λ ein Eigenwert von B und $v \in \mathbb{C}^n$ ein Eigenvektor. Dann gilt:

$$\lambda\langle v, v\rangle = \langle v, \lambda v\rangle = \langle v, Bv\rangle = \langle Bv, v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}\langle v, v\rangle$$

Da $v \neq 0$ ist $\|v\|^2 = \langle v, v\rangle \neq 0$. Also $\lambda = \overline{\lambda}$, d.h., $\lambda \in \mathbb{R}$. $\square$

Insbesondere haben also symmetrische reelle Matrizen $A \in \mathbb{R}^{n \times n}$, d.h., $A = A^t$, nur reelle Eigenwerte.

Beispiel.

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

hat das charakteristische Polynom

$$\chi_A(t) = (2-t)^3 - 2(2-t) = (2-t)(t^2 - 4t + 2)$$

und die Eigenwerte 2 und $2 \pm \sqrt{2}$ in $\mathbb{R}^3$.

Ende Teil 3

Hauptachsentransformation

Satz. *Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann existiert eine orthogonale Matrix $S \in \mathrm{SO}(n)$, so dass*

$$SAS^t = SAS^{-1} = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist.

Beweis. Wir zeigen, $\mathbb{R}^n$ besitzt eine Orthonormalbasis aus Eigenvektoren zu A. Sei λ_1 ein Eigenwert von A. Es gilt $\lambda_1 \in \mathbb{R}$. Sei $v \in \mathbb{R}^n$ ein normierter Eigenvektor von A zum Eigenwert λ. Dann ist das orthogonale Komplement

$$H = v^\perp = \{w \in \mathbb{R}^n \mid \langle w, v\rangle = 0\}$$

ein Untervektorraum, und A bildet H in sich ab:

$$\begin{aligned} w \in H \quad &\Rightarrow \quad \langle w, v\rangle = 0 \\ &\Rightarrow \quad \langle Aw, v\rangle = \langle w, Av\rangle = \langle w, \lambda_1 v\rangle = \lambda_1\langle w, v\rangle = 0 \\ &\Rightarrow \quad Aw \in H. \end{aligned}$$

Mit Induktion nach n dürfen wir annehmen, dass H, welcher isometrisch zu $\mathbb{R}^{n-1}$ ist, eine Orthonomalbasis $\{v_2 \ldots, v_n\}$ aus Eigenvektoren von A besitzt. Die Matrix $(v_1\, v_2 \ldots v_n) \in O(n)$ ist orthogonal und deren Transponierte S erfüllt

$$SAS^t = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Um schließlich $S \in \mathrm{SO}(n)$ zu erreichen, können wir gegebenfalls v_1 durch $-v_1$ ersetzen. □

Beispiel.

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

hat die Eigenwerte 2 und $2 \pm \sqrt{2}$ in $\mathbb{R}^3$.

$$\mathrm{Eig}(A, 2) = \ker \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \mathrm{Spann}(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix})$$

Also können wir

$$v_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ -\frac{\sqrt{2}}{2} \end{pmatrix}$$

wählen. Dessen orthogonales Komplement ist

$$H = \mathrm{Spann}(\begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix})$$

und diese Vektoren bilden eine Orthonormalbasis $\mathcal{A}'$ von H. Die Matrixdarstellung von A in der Basis $\mathcal{A} = \{v_1\} \cup \mathcal{A}'$ ist

$$M_{\mathcal{A}}^{\mathcal{A}}(A) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & -\sqrt{2} \\ 0 & -\sqrt{2} & 2 \end{pmatrix}.$$

In der Tat gilt,

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \\ \sqrt{2} \end{pmatrix} = 2 \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} - \sqrt{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

und

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = -\sqrt{2} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Die Eigenwerte von $A' = \begin{pmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 2 \end{pmatrix}$ sind $2 \pm \sqrt{2}$ und die Eigenräume sind

$$\mathrm{Eig}(A', 2 \pm \sqrt{2}) = \mathrm{Spann}(\begin{pmatrix} 1 \\ \mp 1 \end{pmatrix}) = \mathrm{Spann}(\begin{pmatrix} \frac{\sqrt{2}}{2} \\ \mp \frac{\sqrt{2}}{2} \end{pmatrix}))$$

Also ist

$$S^t = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

eine Matrix mit

$$SAS^t = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{pmatrix}$$

Allerdings ist $S^t \notin \mathrm{SO}(3)$. Wir können also eine Spalte mit -1 multiplizieren, oder, wenn wir die Einträge in der Diagonalmatrix der Größe nach sortiert haben wollen, die erste mit der zweiten Spalte vertauschen, um eine Matrix in $\mathrm{SO}(3)$ zu erhalten.

Ende Teil 4