Technische Universität Chemnitz — Sommersemester 23
Fakultät für Mathematik
H. Schumacher, E. Döhrer

Übungen zu **Mathematische Grundlagen der Computergeometrie**

Übungsblatt 6

Die Aufgaben werden nicht in der Übung besprochen, stattdessen wird es eine Musterlösung geben. Wir legen Ihnen nahe, die Aufgaben trotzdem zu versuchen und die Musterlösung zur Selbstkontrolle zu verwenden.

Übungsaufgabe 6.1

Teilaufgabe a:
Es gilt $P_{\text{span}(a)}(v) = \frac{\langle v,a\rangle}{\|a\|^2}a$ und $P_{V^\perp} = 1 - P_V$. Wir setzen nun $V = \text{span}(a)$, setzen die Vektoren ein und erhalten

$$P_{\text{span}(a)}(v) = \frac{2}{3}a = \frac{1}{3}\begin{pmatrix}2\\2\\2\end{pmatrix}$$
$$P_{\text{span}(a)^\perp}(v) = v - \begin{pmatrix}2\\2\\2\end{pmatrix} = \frac{1}{3}\begin{pmatrix}4\\-2\\-2\end{pmatrix}$$

Teilaufgabe b:
Für ein linear unabhängiges System $A = \{a_1, ..., a_k\}$ ist die Matrix des Projektors auf span(A) gegeben durch

$$P_A = \tilde{A}(\tilde{A}^T\tilde{A})^{-1}\tilde{A}^T \text{ mit } \tilde{A} = (a_1, ..., a_k)$$

Unser $\tilde{A} = \begin{pmatrix}1\\1\\1\end{pmatrix}$. Damit vereinfacht sich die Gleichung zu

$$P_A = \frac{1}{\|a\|^2}aa^T = \frac{1}{3}\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$$

Da $P_{A^\perp} = 1 - P_A$ erhalten wir

$$P_2 = 1 - P_1 = \frac{1}{3}\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix}$$

Wobei P_1 der Projektor auf $\text{span}(A)$ und P_2 das orthogonale Komplement des Spans.
Wir rechnen nun

$$P_1^2 = \frac{1}{9}\begin{pmatrix}3&3&3\\3&3&3\\3&3&3\end{pmatrix} = \frac{1}{3}\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix} = P_1$$
$$P_2^2 = (1-P_1)^2 = 1 + P_1^2 - 2P_1 = 1 + P_1 - 2P_1 = 1 - P_1 = P_2$$

Da P_1 der Projektor auf die Gerade $\text{spam}(a)$ gilt $\text{rang}(P_1) = \dim(image(A)) = \dim(\text{span}(a)) = 1$ und damit gilt automatisch $\det(P_1) = 0$.
Für P_2 verwenden wir nun $\text{rang}(P_2) = 3 - \dim(\text{Ker}(P_2)) = 3 - 1 = 2$, wodurch automatisch $\det(P_2) = 0$ gilt. In der vorherigen Formel haben wir verwendet, dass $\text{Kern}(P_{U^\perp}) = U$.

Teilaufgabe c:
Eine Möglichkeit ist natürlich die Eigenschaft einfach nachzurechnen. Allerdings gibt es einen recht schnellen strukturellen Beweis:
Es sei $w \in \mathbb{R}^3$ beliebig. Wir berechnen

$$\begin{aligned}
\langle P_1(w), P_2(w)\rangle &= \langle w - P_1(w), P_1(w)\rangle = \langle w, P_1(w)\rangle - \|P_1(w)\|^2 \\
&= \frac{1}{\|a\|^2}\langle a, w\rangle \langle w, a\rangle - \frac{\langle a, w\rangle^2}{\|a\|^4}\|a\|^2 = \frac{\langle a, w\rangle^2}{\|a\|^2} - \frac{\langle a, w\rangle^2}{\|a\|^2} \\
&= 0
\end{aligned}$$

Projiziert man auf Unterräume mit Dimension >1, muss man eine orthogonale Basis wählen und Pythagoras nutzen.

Übungsaufgabe 6.2

Gegeben sei eine Ebene E in $\mathbb{R}^3$ durch $\{3x_1 - 2x_2 + x_3 = 0\}$. Indem wir erst $x_3 = 0$ setzen und anschließend $x_1 = 0$ setzen, erhalten wir 2 Lösungen der Gleichung

$$v_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \text{ und } v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Über Gram-Schmidt bestimmen wir nun einen Vektor w aus $\text{span}(v_1, v_2)$ der orthogonal zu v_1 ist:

$$w = v_2 - \frac{\langle v_1, v_2 \rangle}{\|v_1\|^2} v_1 = v_2 - \frac{3}{\sqrt{13}} v_1 = \frac{1}{\sqrt{13}} \begin{pmatrix} -6 \\ \sqrt{13} - 9 \\ 2\sqrt{13} \end{pmatrix}$$

Die ONB ist nun gegeben durch $\{\frac{1}{\sqrt{13}} v_1, \frac{1}{\|w\|} w\}$.
Wir verfahren nun wie in Aufgabe 1 um die Darstellungsmatrix zu erhalten. Als System A wählen wir nun $A = \{v_1, v_2\}$.
Wir berechnen

$$\tilde{A}^T A = \begin{pmatrix} 13 & 3 \\ 3 & 5 \end{pmatrix} \text{ und somit } (\tilde{A}^T A)^{-1} = \frac{1}{56} \begin{pmatrix} 5 & -3 \\ -3 & 13 \end{pmatrix}$$

Damit folgt nun

$$P = \frac{1}{56} \tilde{A} \begin{pmatrix} 10 & 12 & -6 \\ -6 & 4 & 26 \end{pmatrix} = \frac{1}{56} \begin{pmatrix} 20 & 24 & -12 \\ 24 & 32 & 8 \\ -12 & 8 & 52 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 5 & 6 & -3 \\ 6 & 8 & 2 \\ -3 & 2 & 18 \end{pmatrix}$$

Wir wissen bereits, dass $\text{Ker(P)} = E^\perp = \text{span}(v_1 \times v_2) = \text{span}(\begin{pmatrix} 6 \\ -4 \\ 2 \end{pmatrix})$.

Nun bestimmen wir den Abstand des Punkte $v = \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix}$ zur Ebene E. Dieser ist gegeben durch

$$d(v, E) = \frac{|\langle v_1 \times v_2, v \rangle|}{\|v_1 \times v_2\|} = \frac{10}{\sqrt{56}}.$$

Übungsaufgabe 6.3

Es seien $n, k \in \mathbb{N}$ mit $k \leq n$ und $A = \{u_1, ..., u_k\}$ ein System von k linear unbabhängigen Vektoren des $\mathbb{R}^n$.

(a) Geben Sie die Anzahl der Rechenschritte in groß-O-Notation an, die Sie benötigen um einen Vektor x auf span(A) via Gram-Schmidt zu projizieren.

(b) Aus der Vorlesung ist eine Formel für die Darstellungsmatrix der Orthogonalprojektion auf span(A) bekannt. Wie viele Rechenschritte werden in dieser Methode benötigt?

(c) Vergleichen Sie Ihre Ergebnisse. Es Sei k=1. Entscheiden Sie, für welche n die Projektion über Gram-Schmidt effizienter ist.

Teilaufgabe a:

Zunächst ermitteln wir mit dem Gram-Schmidt-Verfahren eine Orthonormalbasis $e_1, \ldots, e_k$ von span$(u_1, ..., u_k)$. Der Orthoprojektor P auf span$(u_1, ..., u_k)$ ist dann durch

$$P = e_1\, e_1^T + e_2\, e_2^T + \cdots + e_k\, e_k^T$$

gegeben.

Für das Gram-Schmidt-Verfahren benötigen wir folgende Operationen:

- Berechnen von $\frac{1}{|v|} = \frac{1}{\sqrt{v^T v}}$ für einen Vektor $v \in \mathbb{R}^n$. Das Skalarprodukt $v^T v$ benötigt n Multiplikation und n Addition. Mit *fused multiply-add* (FMA) sind das n arithmetische Operationen. Zusätzlich benötigen wir eine Division und eine Quadratwurzel oder alternativ eine reziproke Quadratwurzel.

- Multiplizieren eines Vektors $y \in \mathbb{R}^n$ mit einem Skalar $\alpha \in \mathbb{R}$, d.h., $y \leftarrow \alpha\, y$. Dies schlägt mit n Multiplikationen zu Buche.

- Für $\alpha \in \mathbb{R}$ und x, $y \in \mathbb{R}^n$ die Operation. $y \leftarrow y - \alpha\, x$ Mit FMA sind dies auch n arithmetische Operationen.

Die Orthonormalbasis wird wie folgt bestimmt:

Schritt 1:

$$v_1 \leftarrow u_1.$$

Führt man dies in-place aus, so kostet dies nichts.

$$e_1 \leftarrow \frac{1}{|v_1|}\, v_1.$$

Dies sind $2\, n$ arithmetische Operation und eine reziproke Wurzel.

Schritt 2:

$$v_2 \leftarrow u_2 - e_1\, (e_1^T u_2)$$

Dies sind $2\,n$ arithmetische Operation: n für $\alpha_1 = -e_1^T e_2$ und n für $e_2 \leftarrow \alpha_1\, e_1 + e_2$.

$$e_2 \leftarrow \frac{1}{|v_2|}\, v_2.$$

Noch einmal $2\,n$ arithmetische Operation und eine reziproke Wurzel.

Schritt 3:

$$\begin{aligned} v_3 &\leftarrow u_3 - (e_1^T u_3)\, e_1 \\ v_3 &\leftarrow v_3 - (e_2^T v_3)\, e_2 \end{aligned}$$

Dies sind $2 \cdot (2\,n)$ arithmetische Operation; je $2\,n$ für jede Zeile.

$$e_3 \leftarrow \frac{1}{|v_3|}\, v_3.$$

Noch einmal $2\,n$ arithmetische Operation und eine reziproke Wurzel.
Langsam bildet sich ein System heraus:
Schritt $m <= k$:

$$\begin{aligned} v_m &\leftarrow v_m - (e_1^T u_m)\, e_1 \\ v_m &\leftarrow v_m - (e_2^T v_m)\, e_2 \\ \vdots \quad & \quad \vdots \\ v_m &\leftarrow v_m - e_{m-1}\, (e_{m-1}^T v_m) \end{aligned}$$

Dies sind $(m-1) \cdot (2\,n)$ arithmetische Operation; je $2\,n$ für jede Zeile.

$$e_m \leftarrow \frac{1}{|v_m|}\, v_m.$$

Insgesamt brauchen wir k Schritte. Der Rechenaufwand beträgt

$$\underbrace{\sum_{m=1}^{k} 2\,n\,(m-1)}_{\text{Subtraktion der “alten” } e_i} \quad + \quad \underbrace{2\,n\,k}_{\text{Norm und Skalierung}} \quad = n\,k\,(k-1) + 2\,n\,k = n\,k\,(k+1)$$

arithmetische Operation plus 2 k reziproke Wurzeln. Würden wir die Matrix $P = \sum_{m=1}^{k} e_m\, e_m^T$ aufstellen, so würde dies $n^2\,k$ arithmetische Operation kosten und die Matrix-Vektor-Multiplokation schlüge noch einmal mit n^2 arithmetischen Operation zu Buche. Tatsächlich können wir aber $y = Px$ stattdessen wie folgt bekommen:

$$\begin{aligned} y &\leftarrow x - (e_1^T x)\, e_1 \\ y &\leftarrow y - (e_2^T x)\, e_2 \\ \vdots \quad & \quad \vdots \\ y &\leftarrow y - (e_k^T k)\, e_k \end{aligned}$$

Wir wissen bereits, dass hier jeder Schritt $2\,n$ FMA-Operationen kostet. Das heißt, wir können $P\,x$ auch in $2\,n\,k$ Operatonen bekommen. Für große n und k sehr viel kleiner als

n ist das sehr viel günstiger als $n^2\,(k+1)$. Insgesamt erhalten wir als Komplexität der Projektion via Gram-Schmidt:

$$n\,k\,(k+3)\,\text{FMA} + k\,\text{reziproke Wurzeln} \in O(n\,k^2).$$

Teilaufgabe b:
Alternativ haben wir die Formel

$$P\,x = A\,(AA^T)^{-1}A^T\,x.$$

Es gibt nun etliche Möglichkeiten, dies durchzuführen. In der Regel wird es nicht ratsam sein, die Matrix P überhaupt aufzustellen. Stattdessen beobachten wir, dass AA^T symmetrisch und positive-definite ist und wir deshalb eine Cholesky-Zerlegung $AA^T = L^T L$ mit einer unteren Dreiecksmatrix L finden können. Dann gilt $(AA^T)^{-1} = L^{-1}\,(L^T)^{-1}$.

$$\begin{aligned}
z &\leftarrow A^T\,x \\
B &\leftarrow AA^T \\
L &\leftarrow \operatorname{chol}(B) \qquad && \text{Cholesky-Zerlegung } B = L^T L \\
z &\leftarrow (L^T)^{-1} z \\
z &\leftarrow L^{-1} z \\
y &\leftarrow A\,z
\end{aligned}$$

Die Matrix-Vektor-Multiplikation der ersten Zeile kostet $k\,n$ Rechenoperationen.
Die Matrix-Matrix-Multiplikation der zweiten Zeile kostet $n\,k^2$ Rechenoperationen.
Aus der Literatur erhält man für die Cholesky-Zerlegung Kosten von $1/6\,(k+2)\,(k+1)\,k$ FMA-Operationen und k reziproken Quadratwurzeln.
Weil L und L^T Dreiecksmatrizen sind, kann man die Operation ihrer Inversen jeweils durch sogenanntes Rückeinsetzen durch $\frac{1}{2}\,k\,(k-1)$ FMA-Operation und k Divisionen durchführen.
Die letzte Matrix-Vektor-Multiplikation kostet noch einmal $n\,k$ FMA-Operationen. Das sind insgesamt

$$\tfrac{1}{6}\,(k+2)\,(k+1)\,k + k\,(k-1) + n\,k^2 + 2\,n\,k$$

FMA-Operationen und je k Divisionen und k reziproke Quadratwurzeln.

Bemerkung: Tatsächlich kann man sich die Quadratwurzeln bei Gram-Schmidt-Verfahren sparen, indem man die Zahlen $\beta_i \leftarrow \frac{1}{v_i^T v_i}$ speichert und statt $y \leftarrow y - e_i\,(e_i^T\,y)$ immer $y \leftarrow \beta_i\,(v_i\,v_i^T)$ anwendet. Darüber hinaus spart dies auch etliche Skalierungsoperationen. Aber auch die Cholesky-Zerlegung und das Rückeinsetzen kann man im Prinzip ohne Quadratwurzeln umsetzen.

Teilaufgabe c:
Wegen $k \leq n$ ist die zweite Methode ein wenig teurer als das Gram-Schmidt-Verfahren. Obwohl die Formel für die zweite Methode eingängiger ist (und sich daher für theoretische Überlegungen anbietet), wird man in der Praxis für große n eher das Gram-Schmidt-Verfahren bevorzugen, da es stabiler ist, d.h., weil es sich deutlich robuster verhält falls $u_1, \ldots, u_k$ *beinahe* linear abhängig sind.[1]

[1] Wenn die Vektoren $e_1, \ldots, e_k$ aufgrund von Rundungsfehlern nicht Orthonormal sein sollte, so wendet man darauf einfach noch einmal das Gram-Schmidt-Verfahren an. Man sagt übrigens, zweimal Gram-Schmidt sei stets genug für double precision.

Im Spezialfall $k = 1$ brauchen wir nach diesen Formeln $4\,n$ FMA-Operationen für Gram-Schmidt und nur $3\,n+1$ FMA-Operationen. Beide bräuchten noch je eine reziproke Quadratwurzel. Die kann man sich hier aber sparen, indem man einfach

$$\alpha = -\frac{u_1^T\,x}{u_1^T\,u_1} \quad \text{und} \quad y = x + \alpha\,u_1.$$

verwendet. Dies kommt mit $3\,n$ FMAs und einer Division aus. Ein Cholesky-Schritt ohne Normierung läuft auf dasselbe hinaus.

Übungsaufgabe 6.4

Teilaufgabe a:
Wir zeigen zu erst, dass es sich bei b_A um eine symmetrische, bilineare Abbildung handelt.
Seien $u_1, u_2, v \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}$:

$$b_a(\alpha u_1+\beta u_2, v) = (\alpha u_1+\beta u_2)^T Av = (\alpha u_1^T+\beta u_2^T)Av = \alpha u_1^T Av+\beta u_2^T Av = \alpha b_A(u_1, v)+\beta b_A(u_2, v)$$

Da A symmetrisch ist, gilt

$$b_A(u_1, v) = b_A(u_1, v)^T = (u_1^T Av)^T = v^T A^T u_1 = v^T Au_1 = b_A(v, u_1)$$

Die beiden oberen Zeilen implizieren auch die Linearität im zweiten Argument. Wir müssen somit nur noch die positive Definitheit zeigen. Sei $v \in \mathbb{R}^n$ beliebig.
Wegen der positiven Definitheit von A gilt $b_A(v, v) = v^T Av \geq 0$.
Sei nun $v \in \mathbb{R}^n$ sodass $b_A(v, v) = 0$ Es gilt somit $v^T Av = 0$. Da A aber positiv definit, gilt somit $v = 0$. Dies impliziert die positive Definitheit der Bilinearform.

Teilaufgabe b:
Wir prüfen als erstes ob v_1, v_2 orthogonal bzgl. b_A sind

$$v_1^T Av_2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} \right\rangle = 0$$

Somit sind v_1, v_2 orthogonal bzgl. b_A.
Wir prüfen nun die Normiertheit der Vektoren:

$$\|v_1\|_{b_A} = \sqrt{v_1^T Av_1} = \sqrt{2} \text{ und } \|v_2\|_{b_A} = \sqrt{v_2^T Av_2} = \sqrt{3}$$

Somit sind v_1, v_2 orthogonal aber nicht orthonormal bzgl. b_A.
Wir berechnen die Projektion nun via Gram-Schmidt.

$$\begin{aligned} P^A_{\mathrm{span}(v_1,v_2)}(w) &= \frac{b_A(v_1, w)}{\|v_1\|^2_{b_A}} v_1 + \frac{b_A(w, v_2)}{\|v_2\|^2_{b_A}} v_2 \\ &= \frac{4}{2} v_1 + \frac{4}{3} v_2 = \frac{1}{3} \begin{pmatrix} 6 \\ 4 \\ 0 \end{pmatrix} \end{aligned}$$