Übung 5: Newton-Verfahren und -Projektionen

Um die Inverse einer Matrix $A \in \mathbb{R}^{n,n}$ zu berechnen kann man eine LR-Zerlegung von A

$$A = LR \qquad - \mathcal{O}(n^3)$$

berechnen und dann die linearen Gleichungssysteme

i-ter Einheitsvektor

$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$

$$(*) \qquad Ax_i = e_i \qquad - \mathcal{O}(n^2) \cdot n = \mathcal{O}(n^3)$$

für $i = 1, \dots, n$ lösen. Dann ist $X := [x_1, \dots, x_n] \in \mathbb{R}^{n,n}$ wegen

$$AX = [Ax_1, \dots, Ax_n] = [e_1, \dots, e_n] = I$$

die Inverse von A.

<u>In</u> Übung 2 wurde gezeigt, dass die Funktion $f_b : GL(n) \to \mathbb{R}^n$ gegeben durch

$$f_b(A) := A^{-1} b$$

die Ableitung

$$Df_b(A)[E] = -A^{-1} E A^{-1} b$$

hat.

Ähnlich wie bei $(*)$ hat daher die

Funktion $g: GL(n) \to \mathbb{R}^{n,n}$ gegeben durch

$$g(A) := A^{-1}(I) = [A^{-1}e_1, \dots, A^{-1}e_n]$$
$$= [f_{e_1}(A), \dots, f_{e_n}(A)]$$

die Ableitung

$$Dg(A)[E] = [Df_{e_1}(A), \dots, Df_{e_n}(A)]$$
$$= [-A^{-1}EA^{-1}e_1, \dots, -A^{-1}EA^{-1}e_n]$$
$$= -A^{-1}EA^{-1}(I) = -A^{-1}EA^{-1}.$$

Da Konstanten für die Ableitung keine Rolle spielen hat $G: GL(n) \to \mathbb{R}^{n,n}$

gegeben durch $G(X) = X^{-1} - A$

die Ableitung

$$DG(X)[H] = -X^{-1}HX^{-1}.$$

Außerdem ist $G(X^*) = 0$ genau dann wenn X^* die Inverse von A ist.

Um solch ein X^* zu berechnen kann man versuchen das Newton-Verfahren zu benutzen.

Da für $X \in GL(n)$ gilt

$$K = DG(X)[H] = -X^{-1} H X^{-1}$$

$$\Leftrightarrow \quad -XKX = H$$

ist die lineare Abbildung $DG(X)$ überall in $GL(n)$ invertierbar mit

$$\left(DG(X)\right)^{-1}[K] = -XKX.$$

Das Newton-Verfahren lautet somit

$$\begin{aligned} X_{k+1} &= X_k - \left(DG(X_k)\right)^{-1}[G(X_k)] \\ &= X_k + X_k \left(G(X_k)\right) X_k \\ &= X_k + X_k \left(X_k^{-1} - A\right) X_k \\ &= 2X_k - X_k A X_k. \end{aligned}$$

(Verfahren von Schulz)

Die Kosten für eine Iteration sind somit

$$O(n^3)$$

$\overset{\text{i.a.}}{\Rightarrow}$ nicht für die Praxis geeignet.

Projektionen

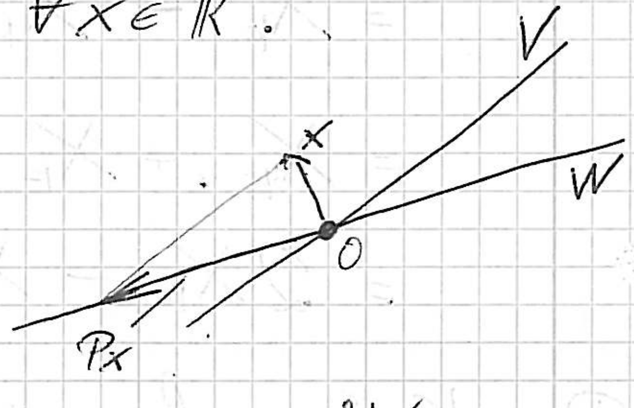
Definition: Sei $P \in \mathbb{R}^{n,n}$

(i) P heißt <u>Projektor auf</u> den Teilraum $W \subseteq \mathbb{R}^n$, falls

$$Pw = w \qquad \forall w \in W, \text{ und}$$
$$Px \in W \qquad \forall x \in \mathbb{R}^n.$$

(ii) P heißt <u>Projektor entlang</u> dem Teilraum $V \subseteq \mathbb{R}^n$, falls

$$Pv = 0 \qquad \forall v \in V \text{ und}$$
$$(x - Px) \in V \qquad \forall x \in \mathbb{R}^n.$$



(iii) P heißt <u>orthogonaler Projektor</u> auf W, wenn es einen Teilraum ~~gibt~~ $W \subseteq \mathbb{R}^n$ gibt, so dass P Projektor auf W entlang $W^\perp$ ist.

Lemma: Es sei $P \in \mathbb{R}^{n,n}$, $V, W \leq \mathbb{R}^n$ gewisse Teilräume und $T \in GL(n)$.
Man setze $\tilde{P} := TPT^{-1}$.

Dann ist

(i) P Projektor auf W genau dann wenn $\tilde{P}$ Projektor auf TW ist.

(ii) P Projektor entlang V genau dann wenn $\tilde{P}$ Projektor entlang TV ist.

Beweis:

(i) $[P$ ist Projektor auf $W]$

$$\Leftrightarrow \left[\cancel{TP}\, w = w \quad \forall w \in W \wedge Px \in W \quad \forall x \in \mathbb{R}^n\right]$$

$$\Leftrightarrow \left[TPT^{-1}Tw = Tw \quad \forall Tw \in TW \wedge TPT^{-1}Tx \in TW \quad \forall Tx \in T\mathbb{R}^n = \mathbb{R}^n\right]$$

$$\Leftrightarrow \left[\tilde{P}\tilde{w} = \tilde{w} \quad \forall \tilde{w} \in TW \wedge \tilde{P}\tilde{x} \in TW \quad \forall \tilde{x} \in \mathbb{R}^n\right]$$

ii) geht analog. $\blacksquare$

Satz: Es sei $P \in \mathbb{R}^{n,n}$. Dann sind äquivalent:

a) $\exists W \subseteq \mathbb{R}^n$ Teilraum, so dass P Projektor auf W ist.

b) $\exists V \subseteq \mathbb{R}^n$ Teilraum, so dass P Projektor entlang V ist.

c) Es ist $PP = P$.

Beweis: (a) $\Rightarrow$ c)) & (b) $\Rightarrow$ c)) sind HA.

c) $\Rightarrow$ a), b): Es sei $T^{-1}PT = J$ die Jordan-Form von P. Dann ist

$$JJ = T^{-1}P\underbrace{TT^{-1}}_{=I}PT = T^{-1}\underbrace{PP}_{=P}T = J.$$

$\Rightarrow$ J hat nur 1 oder 0 auf der Diagonalen

$\Rightarrow$ J ist Diagonalmatrix. O.B.d.A. sei $J = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$,

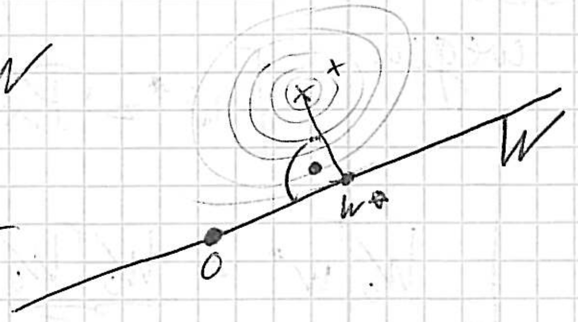
d.h. J sei Projektor auf $\begin{bmatrix} I \\ 0 \end{bmatrix}$ entlang $\begin{bmatrix} 0 \\ I \end{bmatrix}$.

Partitioniere $T =: [X \; Y]$. Dann ist nach obigem Lemma $P = TJT^{-1}$ Projektion auf

$T \operatorname{span}\begin{bmatrix} I \\ 0 \end{bmatrix} = \operatorname{span}(X) =: W$ entlang

$T \operatorname{span}\begin{bmatrix} 0 \\ I \end{bmatrix} = \operatorname{span}(Y) =: V$. $\blacksquare$

Definition: Sei $W \subseteq \mathbb{R}^n$ ein Teilraum und $x \in \mathbb{R}^n$. Dann heißt $w^* \in W$ ***Element bester Approximation*** wenn



$$\|x - w^*\|_2 = \min_{w \in W} \|x - w\|_2 .$$

Da $0 \in W$, ist $\min_{w \in W} \|x - w\|_2 < \|x - 0\|_2 = \|x\|_2$, womit wegen der Kompaktheit von

$$W \cap \{y \in \mathbb{R}^n \mid \|y\|_2 \leq \|x\|_2\}$$

die Existenz eines w^* folgt.

Satz: Sei $W \subseteq \mathbb{R}^n$ ein Teilraum und $x \in \mathbb{R}^n$. Sei $W_0 \in \mathbb{R}^{n,m}$ eine Matrix deren Spalten eine Orthonormalbasis von W bilden. ~~Es sei~~ Dann ist

$$P := W_0 W_0^T$$

(der) orthogonal Projektor auf W. Es sind äquivalent: (für $w^* \in W$)

1.) w^* ist Element bester Approximation

2.) $x - w^* \perp W$

3.) $w^* = Px$

Beweis: Die Äquivalenz von 2.) & 3.) folgt wegen

$$[w^* = Px = W_0 W_0^T x] \overset{W_0^T}{\Longleftrightarrow}$$

$$\left[W_0^T w^* - \underbrace{W_0^T W_0}_{=I} W_0^T x = 0 \right]$$

$$\Leftrightarrow \left[W_0^T (w^* - x) = 0 \right]$$

$$\Leftrightarrow \left[w^* - x \perp W \right].$$

$w^* = W_0 \alpha$
$\Rightarrow \underbrace{W_0^T W_0 \alpha}_{=\alpha} = W_0^T x$
$\Rightarrow w^* = W_0 \alpha = W_0 W_0^T x = Px$

Für die Äquivalenz von 1.) & 2.) sieht man „wie oben", dass

$$\bar{w} := Px$$

die Bedingung $(\bar{w} - x) \perp W$ erfüllt. Damit gilt für alle $w \in W$, dass

$$\|x - w\|_2^2 = \|(x - \bar{w}) - (w - \bar{w})\|_2^2$$

$$= \|x - \bar{w}\|_2^2 - 2 \underbrace{\langle x - \bar{w}, w - \bar{w} \rangle}_{=0,\ \text{da } w - \bar{w} \in W} + \|w - \bar{w}\|_2^2$$

$$= \|x - \bar{w}\|_2^2 + \|w - \bar{w}\|_2^2 \geq \|x - \bar{w}\|_2^2,$$

wobei Gleichheit gilt genau dann wenn $w = \bar{w} = w^*$.

$\blacksquare$