Beispiel 1, Übungsblatt 4

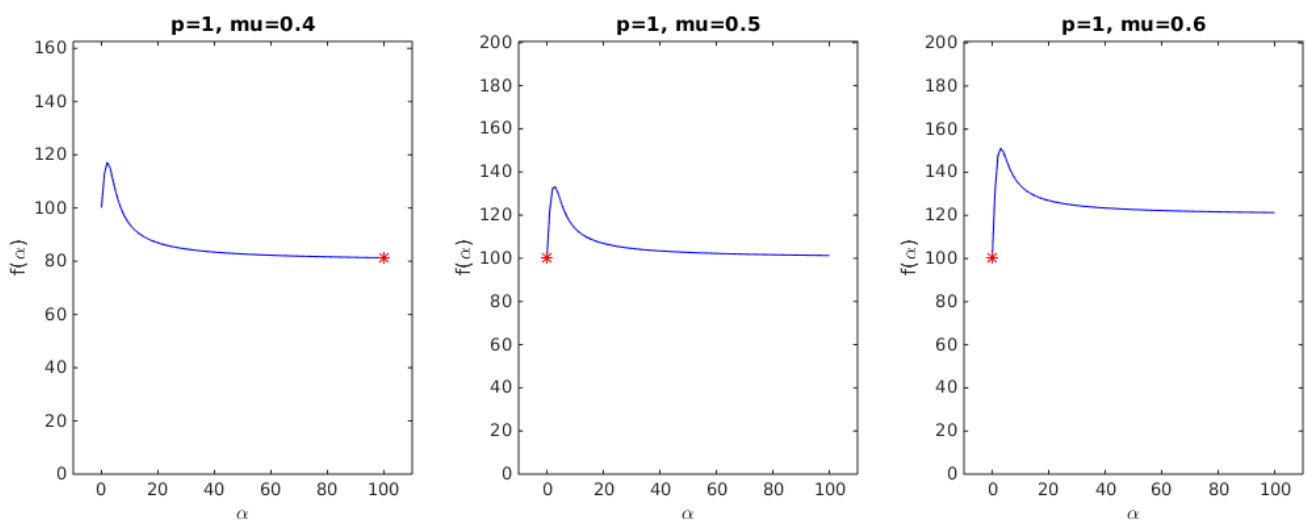
Für $n = 100$ seien $h = 1/n$, $\boldsymbol{x} = \{(i - \frac{1}{2})h\}_{i=1}^{n}$ und $\boldsymbol{v} = \operatorname{sign}(\boldsymbol{x} - \frac{1}{2})$ definiert. Für ausgewählte $\mu \in (0, \infty)$ und $p \in [1, \infty]$ sei $J : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$J(\boldsymbol{u}) = \|\boldsymbol{u} - \boldsymbol{v}\|_p + \mu \|Q\boldsymbol{u}\|_p \quad \text{wobei} \quad Q = \frac{1}{h} \begin{bmatrix} -1 & +1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & +1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}.$$

Die Funktion $f(\alpha) = J(\boldsymbol{u}(\alpha))$ mit $\boldsymbol{u}(\alpha) = \tanh(\alpha(\boldsymbol{x} - \frac{1}{2}))$ und $\alpha \in [0, \infty)$ soll minimiert werden. Kreuzen Sie bei den wahren Behauptungen an.

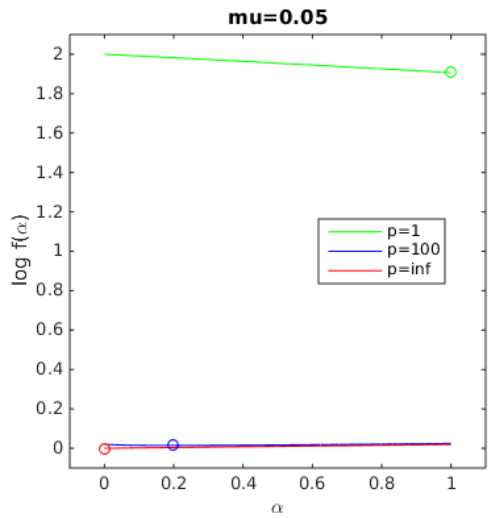
1. Für $p = 1$ und μ ausreichend groß wird f an der Stelle $\alpha^\star = 0$ global minimiert, aber für μ ausreichend klein gibt es kein minimierendes $\alpha \in [0, \infty)$.
2. Für ein gewisses $\mu \in (0, \infty)$ gibt es ein $p \in (1, \infty)$, wobei f an einer Stelle $\alpha_p^\star \in (0, \infty)$ global minimiert wird, während $\alpha_1^\star = \infty$ und $\alpha_\infty^\star = 0$ gelten, d.h. $\inf_\alpha f(\alpha) = \lim_{\alpha\to\infty} f(\alpha)$ für $p = 1$ und $\min_\alpha f(\alpha) = f(0)$ für $p = \infty$.
3. Für $p = \infty$ wird f für jedes $\mu \in (0, \infty)$ an der Stelle $\alpha = 0$ global minimiert.
4. Für beliebige $\mu \in (0, \infty)$ und $p \in (1, \infty)$ ist f konvex.

Teil (a): Die Funktion f wird für die Werte $\mu = 0.4$, 0.5 und 0.6 über das Intervall $\alpha \in [0, 100]$ folgendermaßen grafisch dargestellt, wobei das Minimum mit einem roten Sternchen markiert wird.



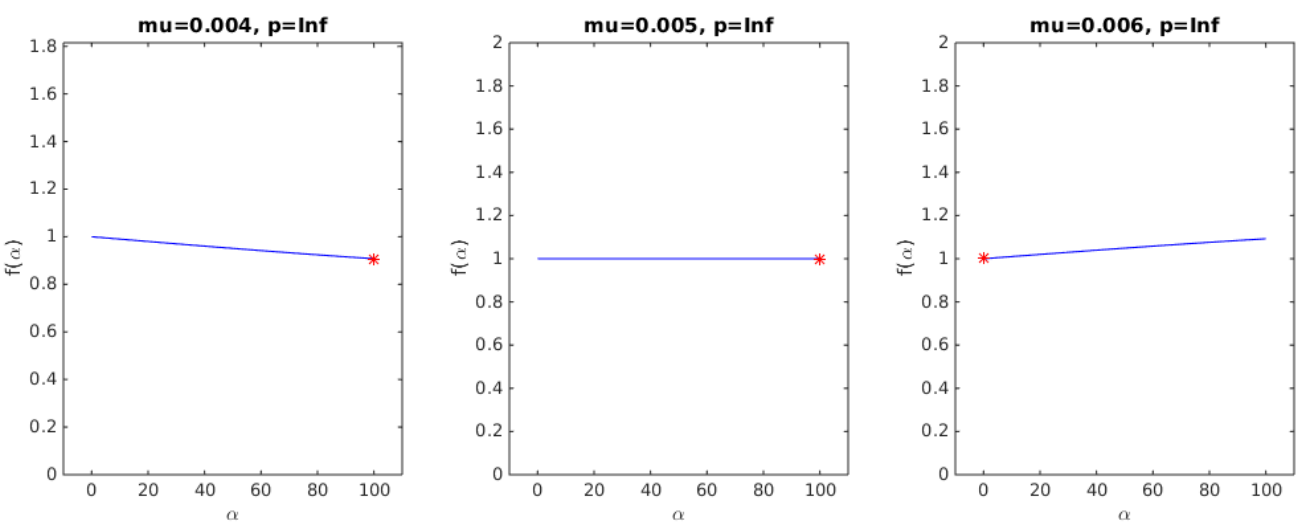
An einer Schwelle $\mu^\star \approx 0.5$ gibt es einen Übergang von einem Infimum bei $\alpha \to \infty$ für $\mu < \mu^\star$ zu einem Minimum bei $\alpha = 0$ für $\mu \geq \mu^\star$.

Teil (b): Aus Teil (a) ist $\mu \approx 0.5$ eine Schwelle für den Fall $p = 1$, und aus Teil (c) ist $\mu \approx 0.005$ eine Schwelle für den Fall $p = \infty$. Für den Wert $\mu = 0.05$ wird die Funktion f für $p = 1, 100, \infty$ über das Intervall $\alpha \in [0, 1]$ folgendermaßen grafisch dargestellt, wobei die jeweiligen Minima $a_p^\star$ mit Kreisen markiert werden.



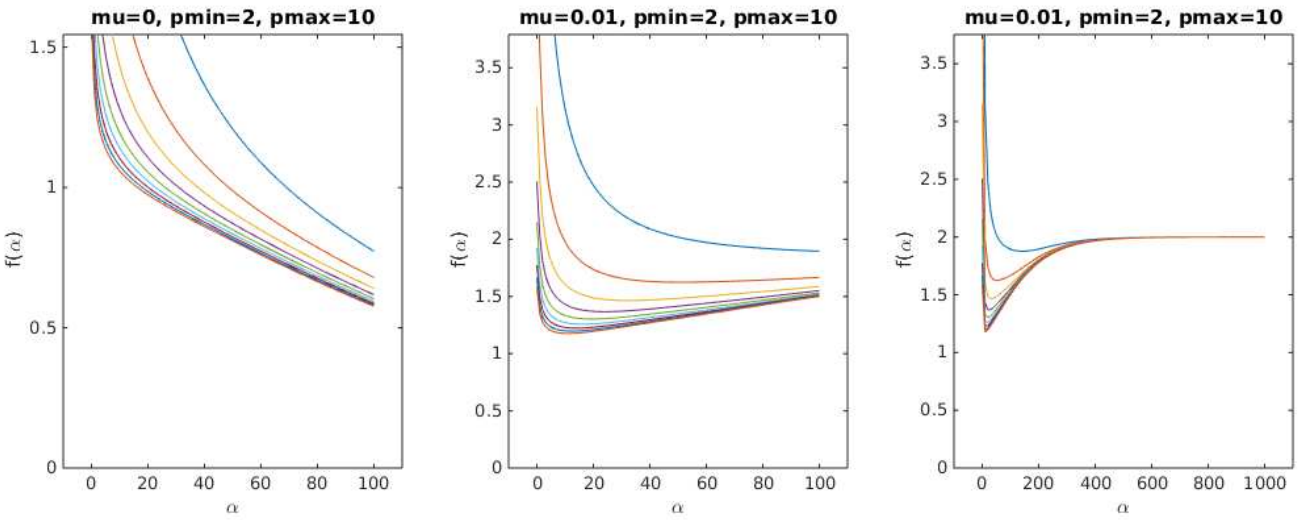
Die grüne Kurve entspricht $p = 1$, die blaue Kurve $p = 100$ und die rote Kurve $p = \infty$. Mit weiteren Details über die Fälle $p = 1$ und $p = \infty$ bei den Teilen (a) bzw. (c) wird die Behauptung mit den Kreisen bestätigt.

Teil (c): Mit $p = \infty$ wird die Funktion f für $\mu = 0.004$, 0.005 und 0.006 über das Intervall $\alpha \in [0, 100]$ folgendermaßen grafisch dargestellt, wobei die jeweiligen Minima mit roten Sternchen markiert werden.



An einer Schwelle $\mu^\star \approx 0.005$ gibt es einen Übergang von einem Infimum bei $\alpha \to \infty$ für $\mu < \mu^\star$ zu einem Minimum bei $\alpha = 0$ für $\mu \geq \mu^\star$.

Teil (d): Mit den Werten $\mu = 0$ und $\mu = 0.01$ wird die Funktion f für $p = 2, \ldots, 10$ über die Intervalle $\alpha \in [0, 100]$ und $\alpha \in [0, 1000]$ folgendermaßen grafisch dargestellt.



Für $\mu = 0$ ist f global konvex. Obwohl für $\mu = 0.01$ die Funktion in der Nähe des Minimums lokal konvex ist, ist sie auf einem größeren Intervall nicht konvex.

3a) Matrixnorm $\|\cdot\|_M$ verträglich mit Vektornorm $\|\cdot\|_V$,

FALSCH falls $\forall x, \forall M$: $\|Ax\|_V \leq \|A\|_M \|x\|_V$

$\|\cdot\|_F$ Frobenius-Norm mit $\|A\|_F := \sqrt{\sum_{i,j=1}^{n,m} |A_{ij}|^2}$

~~xx~~ $\|\cdot\|_F$ verträglich mit $\overset{\sum_{i=1}^{n}|x_i|}{\|\cdot\|_1}$: $\underbrace{\|Ax\|_1}_{\sum_{i,j=1}^{n,m}|a_{ij}x_i|} \leq \|A\|_F \cdot \|x\|_1$

nicht ausreichend,
Gegenbeispiel notwendig

$$\|Ax\|_1 = \sum_{i,j=1}^{n,m} |a_{ij}x_i| \leq \underbrace{\left(\sum_{i,j}^{n,m} |a_{ij}|\right)}_{\underbrace{\sum_{i=1}^{n}}_{\leq \sqrt{n}\left(\sum_{i=1}^{n} 1^2\right)^{1/2}} \cdot \underbrace{\sum_{j=1}^{m}}_{} |a_{ij}|} \cdot \left(\sum_{i=1}^{n}|x_i|\right) \leq \sqrt{nm}\left(\sum_{i,j=1}^{n,m}|a_{ij}|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n}|x_i|\right)$$

$$= \underbrace{\sqrt{nm}}_{\text{für } n,m > 1,\ \sqrt{nm} > 1} \cdot \|A\|_F \cdot \|x\|_1$$

Gegenbeispiel:
A = [1,2,3;2,3,4;3,4,5]; x = [1;2;3];
||Ax||_1 = 60
||A||_F ~= 9.6437
||A||_F ||x||_1 ~= 57.8619 < ||Ax||_1

Matrixnorm folgende Eigenschaften :

(i) $\|A\| \geq 0$ Positivität

(ii) $\|A\| = 0 \Rightarrow A = 0$

(iii) $\|\alpha A\| = |\alpha| \cdot \|A\|$ Homogenität

(iv) $\|A+B\| \leq \|A\| + \|B\|$ Δ-Ungl.

(v) $\|AB\| \leq \|A\| \cdot \|B\|$ Submultiplikativität

3b) $n \in \mathbb{N}$, $n \geq 2$ $\quad \|A\|_{s,t} = \sum_{i,j=1}^{n} |a_{ij}|$, $A = \{a_{ij}\}_{i,j=1}^{n} \in \mathbb{R}^{n \times n}$

WAHR

(i) $\|A\|_{s,t} \geq 0$

(ii) $\|A\|_{s,t} = \sum_{i,j=1}^{n} |a_{ij}| \geq 0 \Rightarrow \forall i,j \text{ mit } |a_{ij}| = 0 \Rightarrow A = 0$

kann nur 0, wenn alle Summanden gleich 0

(iii) $\|\alpha A\|_{s,t} = \sum_{i,j=1}^{n} |\alpha a_{ij}| = |\alpha| \cdot \sum_{i,j=1}^{n} |a_{ij}| = |\alpha| \cdot \|A\|_{s,t}$

(iv) $\|A+B\|_{s,t} = \sum_{i,j=1}^{n} |a_{ij} + b_{ij}| \leq \sum_{i,j=1}^{n} |a_{ij}| + \sum_{i,j=1}^{n} |b_{ij}| = \|A\|_{s,t} + \|B\|_{s,t}$

(v) $\|A \cdot B\|_{s,t} = \sum_{i,j=1}^{n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \leq \sum_{i,j,k=1}^{n} |a_{ik}| \cdot |b_{kj}| = \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ik}| \cdot \underbrace{|b_{kj}|}_{\leq \sum_{l=1}^{n} |b_{lj}|}$

$\leq \left(\sum_{i,k=1}^{n} |a_{ik}| \right) \cdot \left(\sum_{j,l=1}^{n} |b_{lj}| \right) = \|A\|_{s,t} \cdot \|B\|_{s,t}$

3c) $n \in \mathbb{N}$, $n \geq 2$ $\quad \|A\|_{m,t} = \max_{1 \leq i,j \leq n} |a_{ij}|$, $\quad A = \{a_{ij}\}_{i,j=1}^{n} \in \mathbb{R}^{n \times n}$

FALSCH

(i) $\|A\|_{m,t} \geq 0$

(ii) $\|A\|_{m,t} = \max_{1 \leq i,j \leq n} |a_{ij}| \geq 0 \implies \forall i,j \text{ mit } |a_{ij}| = 0 \implies A = 0$

(iii) $\|\alpha A\|_{m,t} = \max_{1 \leq i,j \leq n} |\alpha \cdot a_{ij}| = \max_{1 \leq i,j \leq n} (|\alpha| \cdot |a_{ij}|) = |\alpha| \cdot \max_{1 \leq i,j \leq n} |a_{ij}| = |\alpha| \cdot \|A\|_{m,t}$

(iv) $\|A+B\|_{m,t} = \max_{1 \leq i,j \leq n} |a_{ij} + b_{ij}| \leq \max_{1 \leq i,j \leq n} (|a_{ij}| + |b_{ij}|) \leq \max_{1 \leq i,j \leq n} |a_{ij}| + \max_{1 \leq i,j \leq n} |b_{ij}| = \|A\|_{m,t} + \|B\|_{m,t}$

(v) $\|A \cdot B\|_{m,t} = \max_{1 \leq i,j \leq n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \leq \max_{1 \leq i,j \leq n} \sum_{k=1}^{n} (|a_{ik}| \cdot |b_{kj}|) \leq n \cdot \|A\| \cdot \|B\|$

nicht ausreichend, Gegenbeispiel notwendig:
A = [1,1;0,1], B = [1,0;1,1]
||A*B||=||[2,1;1,1]||=2, ||A||=||B||=1

(i) bis (iv) erfüllt, aber (v) nicht

3d) zz: $\|A\|_\varepsilon < \rho(A) + \varepsilon \quad \forall \varepsilon > 0$

WAHR

mit $A = S \Lambda S^{-1}$, Λ Diagonalmatrix der EW

$$\|A\|_\varepsilon = \| D(\tfrac{1}{\varepsilon}) \underbrace{S^{-1} A S}_{= \Lambda} D(\varepsilon) \|_\infty$$

$$= \| D(\tfrac{1}{\varepsilon}) \Lambda D(\varepsilon) \|_\infty \overset{(*)}{=} \max_{j=1,\dots,n} \sum_{i=1}^{n} |\Lambda_{ij}|$$

da Λ Diagonalmatrix der EW $(*) = \|\Lambda\|_\infty$

da Λ Diagonalmatrix ist $\|\Lambda\|_\infty$ größter Wert von Λ $\triangleq$ größtem EW von A

$\Longrightarrow \|A\|_\varepsilon = \rho(A) \Longrightarrow \|A\|_\varepsilon < \rho(A) + \varepsilon$

codingground SIMPLY EASY CODING | Execute MATLAB/Octave Online (GNU Octave, v4.2.1)

Fork Project

Execute | Share | demo.m | STDIN

```
A=[2,-1,1;2,2,2;-1,-1,2];

b=[-1;4;-5];
x=b;

D=diag(diag(A));
D=inv(D);

L=tril(A,-1);
U=triu(A,+1);
R=L+U;

tol = 10^-8; %rel Fehlertoleranz
itmax = 1000; %max Anzahl Iterationen
k = 1; %Anzahl Iterationen

for y=1:1000

    x1 = D*(b-R*x);
    err = norm(x1-x)/norm(x1);
    x = x1;

    k = k+1;

        if err <= tol
            break
        end

end

disp(x1)
k
err
```

Result

```
$octave -qf --no-window-system demo.m
  -6.8425e+48
   6.8425e+48
  -1.0264e+49
k =  1001
err =  1.6270
warning: function ./demo.m shadows a core library function
```

codingground SIMPLY EASY CODING

Execute MATLAB/Octave Online (GNU Octave, v4.2.1)

Fork

Execute | Share | demo.m | STDIN

```
A=[2,-1,1;2,2,2;-1,-1,2];

b=[-1;4;-5];
x=b;
tol = 10^-5; %rel Fehlertoleranz
k=1;
n=size(A,1);

for k=1:1000
    for i=1:n
        s(i)=A(i,:)*x;


        x1(i)=x(i)-1/A(i,i)*(s(i)-b(i));


            if abs(x1(i)-x(i))<tol;
                break
            end

        x(i)=x1(i);


    end


end

x
```

Result

```
$octave -qf --no-window-system demo.m
x =

   1.0000
   2.0000
  -1.0000

warning: function ./demo.m shadows a core library function
```

codingground SIMPLY EASY CODING

Execute MATLAB/Octave Online (GNU Octave, v4.2.1)

Fork

Execute | Share | demo.m | STDIN

```
A=[1,2,-2;1,1,1;2,2,1];

b=[7;2;5];
x=b;

D=diag(diag(A));
D=inv(D);

L=tril(A,-1);
U=triu(A,+1);
R=L+U;

tol = 10^-8; %rel Fehlertoleranz
itmax = 1000; %max Anzahl Iterationen
k = 1; %Anzahl Iterationen

for y=1:1000

    x1 = D*(b-R*x);
    err = norm(x1-x)/norm(x1);
    x = x1;

    k = k+1;

        if err <= tol
            break
        end

end

disp(x1)
k
err
```

Result

```
$octave -qf --no-window-system demo.m
   1
   2
  -1
k =  4
err = 0
warning: function ./demo.m shadows a core library function
```

codingground SIMPLY EASY CODING

Execute MATLAB/Octave Online (GNU Octave, v4.2.1)

Fork

Execute | Share | demo.m | STDIN

```
A=[1,2,-2;1,1,1;2,2,1];

b=[7;2;5];
x=b;
tol = 10^-5; %rel Fehlertoleranz
k=1;
n=size(A,1);

for k=1:1000
    for i=1:n
        s(i)=A(i,:)*x;


        x1(i)=x(i)-1/A(i,i)*(s(i)-b(i));


            if abs(x1(i)-x(i))<tol;
                break
            end

        x(i)=x1(i);


    end


end

x
```

Result

```
$octave -qf --no-window-system demo.m
x =

   9.6404e+304
  -9.6436e+304
   6.4291e+301

warning: function ./demo.m shadows a core library function
```

Numerische Mathematik 1

Übungsblatt 4

7. Für $n \in \mathbb{N}$, $n \geq 2$, $T \in \mathbb{R}^{n \times n}$ und $c \in \mathbb{R}^n$ sei $x^{(k+1)} = Tx^{(k)} + c$ eine Iteration mit mindestens einem Fixpunkt $x^* \in \mathbb{R}^n$. Kreuzen Sie bei den wahren Behauptungen an.

a) Wenn $\|T\|_\natural < 1$ für eine induzierte Norm $\|\cdot\|_\natural$ gilt, dann ist der Fixpunkt x^* eindeutig.

Lösung. [Wahr]

Aus $x^{(k+1)} = Tx^{(k)} + c$ und $x^* = Tx^* + c$ es ergibt sich $(x^{(k+1)} - x^*) = T(x^{(k)} - x^*)$.

Es folgt:

$$\|x^{(k+1)} - x^*\|_\natural \leq \|T\|_\natural \cdot \|x^{(k)} - x^*\|_\natural \leq \ldots \leq \underbrace{\underbrace{\|T\|_\natural^k}_{<1}}_{\xrightarrow[k\to\infty]{} 0} \cdot \|x^{(0)} - x^*\|_\natural \xrightarrow{k\to\infty} 0$$

Nun, angenommen es existiert eine andere Fixpunkt x_2^* der Iteration $x^{(k+1)} = Tx^{(k)} + c$ (d.h. $x^* \neq x_2^*$)

$$\Rightarrow \|x^* - x_2^*\|_\natural = \|x^* - x^{(k+1)} + x^{(k+1)} - x_2^*\|_\natural \leq \underbrace{\|x^* - x^{(k+1)}\|}_{\xrightarrow[k\to\infty]{} 0} + \underbrace{\|x^{(k+1)} - x_2^*\|}_{\xrightarrow[k\to\infty]{} 0}$$

$$\Rightarrow \boxed{x^* = x_2^*} \quad ↯$$

Also, x^* ist eindeutig. ■

b) Wenn es einen eindeutigen Fixpunkt gibt, dann gilt $\rho(T) < 1$.

Lösung. [Falsch]

Nehme $T = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ und $c = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Dann aus $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

$\Rightarrow \underbrace{\boxed{x = 1,\ y = 1}}_{\text{eindeutig}}$, aber $\rho(T) = 2 > 1$.

c) Wenn $\|T\|_\natural < 1$ für eine induzierte Norm $\|\cdot\|_\natural$ gilt, dann folgt

$$\|x^k - x^*\|_\natural \leq \frac{\|T\|_\natural^k}{1-\|T\|_\natural} \|x^{(1)} - x^{(0)}\|_\natural, \forall k \in \mathbb{N}, \forall x^{(0)} \in \mathbb{R}^n.$$

<u>Lösung.</u> [Wahr]

Es gilt für $m \in \mathbb{N}$

$$\|x^{(k+m)} - x^{(k)}\|_\natural = \|x^{(k+m)} - x^{(k+m-1)} + x^{(k+m-1)} - x^{(k+m-2)} + x^{(k+m-2)} - + \ldots - x^{(k)}\|_\natural$$

$$\leq \|x^{(k+m)} - x^{(k+m-1)}\|_\natural + \|x^{(k+m-1)} - x^{(k+m-2)}\|_\natural + \ldots + \|x^{(k+1)} - x^{k}\|_\natural$$

$$= \sum_{p=k}^{k+m-1} \|x^{(p+1)} - x^{(p)}\|_\natural \leq \sum_{p=k}^{k+m-1} \|T\|_\natural^p \cdot \|x^{(1)} - x^{(0)}\|_\natural =$$

$$= \|x^{(1)} - x^{(0)}\|_\natural \cdot \sum_{p=k}^{k+m-1} \|T\|_\natural^p = \|T\|_\natural^k \cdot \|x^{(1)} - x^{(0)}\|_\natural \cdot \sum_{i=0}^{m-1} \|T\|_\natural^i =$$

$$= \|T\|_\natural^k \cdot \|x^{(1)} - x^{(0)}\|_\natural \cdot \frac{1-\|T\|_\natural^m}{1-\|T\|_\natural} \quad \Big| \lim_{m\to\infty}$$

$$\Rightarrow \|x^{(k)} - x^*\|_\natural = \lim_{m\to\infty} \|x^{(k)} - x^{(k+m)}\|_\natural \leq \underbrace{\left(\lim_{m\to\infty} \frac{1-\|T\|^m}{1-\|T\|_\natural}\right)}_{\to \frac{1}{1-\|T\|}} \cdot \|T\|^k \cdot \|x^{(1)} - x^{(0)}\|$$

$$\Leftrightarrow \boxed{\|x^k - x^*\| \leq \frac{\|T\|^k}{1-\|T\|_\natural} \cdot \|x^{(1)} - x^{(0)}\|}$$

d) Für eine beliebige Iterationsmatrix $T \in \mathbb{R}^{n \times n}$ und eine induzierte Norm $\|\cdot\|_{\square}$ folgt

$$\|x^{(k)} - x^*\|_{\square} = \|x^{(0)} - x^*\|_{\square} \cdot \|T\|_{\square}^{k} \cdot \sum_{m=0}^{k-1} \|T\|_{\square}^{m}, \quad \forall k \in \mathbb{N}, \forall x^{(0)} \in \mathbb{R}^n.$$

<u>Lösung.</u> [Falsch]

Es gilt: $(x^{(k)} - x^*) = T(x^{(k-1)} - x^*)$

$$\|x^{(k)} - x^*\|_{\square} = \|T \cdot (x^{(k-1)} - x^*)\|_{\square} \leq \|T\| \cdot \|x^{(k-1)} - x^*\|_{\square} \leq \ldots \leq \|T\|_{\square}^{k} \cdot \|x^{(0)} - x^*\|_{\square}$$

Nun, betrachte $\displaystyle\sum_{m=0}^{k-1} \|T\|_{\square}^{m} = \frac{1 - \|T\|_{\square}^{k}}{1 - \|T\|_{\square}} \overset{i.A}{\neq} 1$

Für $\|T\|_{\square} > 1$

$$\Rightarrow \quad \|T\|_{\square}^{k} \cdot \|x^{(0)} - x^*\| < \|T\|_{\square}^{k} \cdot \|x^{(0)} - x^*\|_{\square} \cdot \frac{1 - \|T\|_{\square}^{k}}{1 - \|T\|_{\square}}$$

und somit $\|x^{(k)} - x^*\|_{\square} \neq \|x^{(0)} - x^*\|_{\square} \cdot \|T\|_{\square}^{k} \cdot \sum_{m=0}^{k-1} \|T\|_{\square}^{m}$. ■