22.3.2012

Bsp (Arithmetik mit/ohne guard-bit)

Bsp: $(1.0)_{10} - (0.875)_{10} = (0.125)_{10} = \frac{1}{8}$

$\quad\;\; \| \qquad\qquad\quad \|$

$x = (1.00)_2 \cdot 2^0 \quad y = (1.11) \cdot 2^{-1}$

$\Rightarrow$ ohne guard-bit

$$\begin{array}{rl} & 1.00 \cdot 2^0 \\ - & 0.11 \cdot 2^0 \\ \hline & 0.01 \cdot 2^0 \end{array}$$

$\underbrace{0.01 \cdot 2^0}_{=\frac{1}{4}}$

$\Rightarrow$ mit guard-bit

$$\begin{array}{rl} & 1.000 \cdot 2^0 \\ - & 0.111 \cdot 2^0 \\ \hline & 0.001 \cdot 2^0 \end{array}$$

$\underbrace{0.001 \cdot 2^0}_{=\frac{1}{8}}$

absoluter Fehler:

$|(x \ominus y) - (x-y)|$

$= \frac{1}{8}$

kein Fehler $\Rightarrow$ Ergebnis exakt !

relativer Fehler:

$$\frac{|(x \ominus y) - (x-y)|}{(x-y)} = \frac{\frac{1}{8}}{\frac{1}{8}} = 1 = \varepsilon_M$$

Bsp (Rechengesetze)

Summe von x, y, z:

$(x \oplus y) \oplus z = (x+y)(1+\delta_1) \oplus z$

$= [(x+y)(1+\delta_1) + z](1+\delta_1)$ Term vernachlässigen

$= (x+y+z) + (x+y)(\underbrace{\delta_1}_{\leq \varepsilon_M} + \underbrace{\delta_2}_{\leq \varepsilon_M} + \delta_1 \cdot \delta_2) + z \cdot \delta_2 \cdot x \oplus (y \oplus z)$

$= x \oplus (y \oplus z) = x \oplus (y+z) \cdot (1+\delta_3)$

$= [x + (y+z) \cdot (1+\delta_3)] \cdot (1+\delta_4)$

$= (x+y+z) + (y+z)(\delta_3 + \delta_4 + \delta_3 \cdot \delta_4) + x \cdot \delta_4$

$\Rightarrow |(x \oplus y) \oplus z - (x+y+z)| \dot{\leq} 2|x+y| \cdot \varepsilon_M + |z| \cdot \varepsilon_M$

(Der Punkt weist auf vernachlässigte Terme hin)

$$= \underbrace{(2|x+y|+|z|)}_{C_1} \cdot \varepsilon_M$$

$$\Rightarrow |x \oplus (y \oplus z) - (x+y+z)| \leq \underbrace{(2|y+z|+|x|)}_{C_2} \cdot \varepsilon_M$$

Falls $|x| \gg |y|$ und $|x| \gg |z|$

So ist $C_2 < C_1$

Merke

Summation mit betragskleinsten Elementen beginnen. $\left(\sum_{k=1} \frac{1}{k} \right)$

Bsp (Landau - Symbol)

a) $\sin(x) = \mathcal{O}(x)$ für $(x \to \infty)$

denn $\lim_{x \to 0} \left| \frac{\sin(x)}{x} \right| = 1 < \infty$

b) $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$

$$\Rightarrow e^x - \sum_{k=0}^{n} \frac{x^k}{k!} = \sum_{k=n+1}^{\infty} \frac{x^k}{k!} = \mathcal{O}(x^{n+1}) \text{ für } x \to 0$$

denn $\lim_{x \to 0} \frac{\sum_{k=n+1}^{\infty} \frac{x^k}{k!}}{x^{n+1}} = \frac{1}{(n+1)!} < \infty$

c) Abschätzung: Aufwand Rechenoperationen

<u>Bsp</u>: Matrix-Multiplikation $n \in \mathbb{N}$ Problemgröße

$$\underset{(n \times n)}{A} \cdot \underset{(n \times n)}{B} = \underset{(n \times n)}{C}$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

Aufwand um c_{ij} zu berechnen:

n Multiplikationen

$n-1$ Additionen

$\sum 2n-1$ flop

Gesamtaufwand:

$n^2(2n-1) = 2n^3 - n^2$ flop

$= \mathcal{O}(n^3)$ → Kubische Laufzeit

denn $\lim_{n \to \infty} \frac{2n^3 - n^2}{n^3} = \lim_{n \to \infty} 2 - \frac{1}{n} = 2 < \infty$

<u>Bsp (Normen)</u>

$v = (1,2,3)$

a) p-Norm mit $p = 3$

$\|v\|_3 = \sqrt[3]{|1|^3 + |2|^3 + |3|^3} = \sqrt[3]{36}$

in Matlab: v = [1; 2; 3] norm(v,3)

b) 1-Norm

$\|v\|_1 = |1| + |2| + |3| = 6$

c) 2-Norm (Euklidische Norm)

$\|v\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$

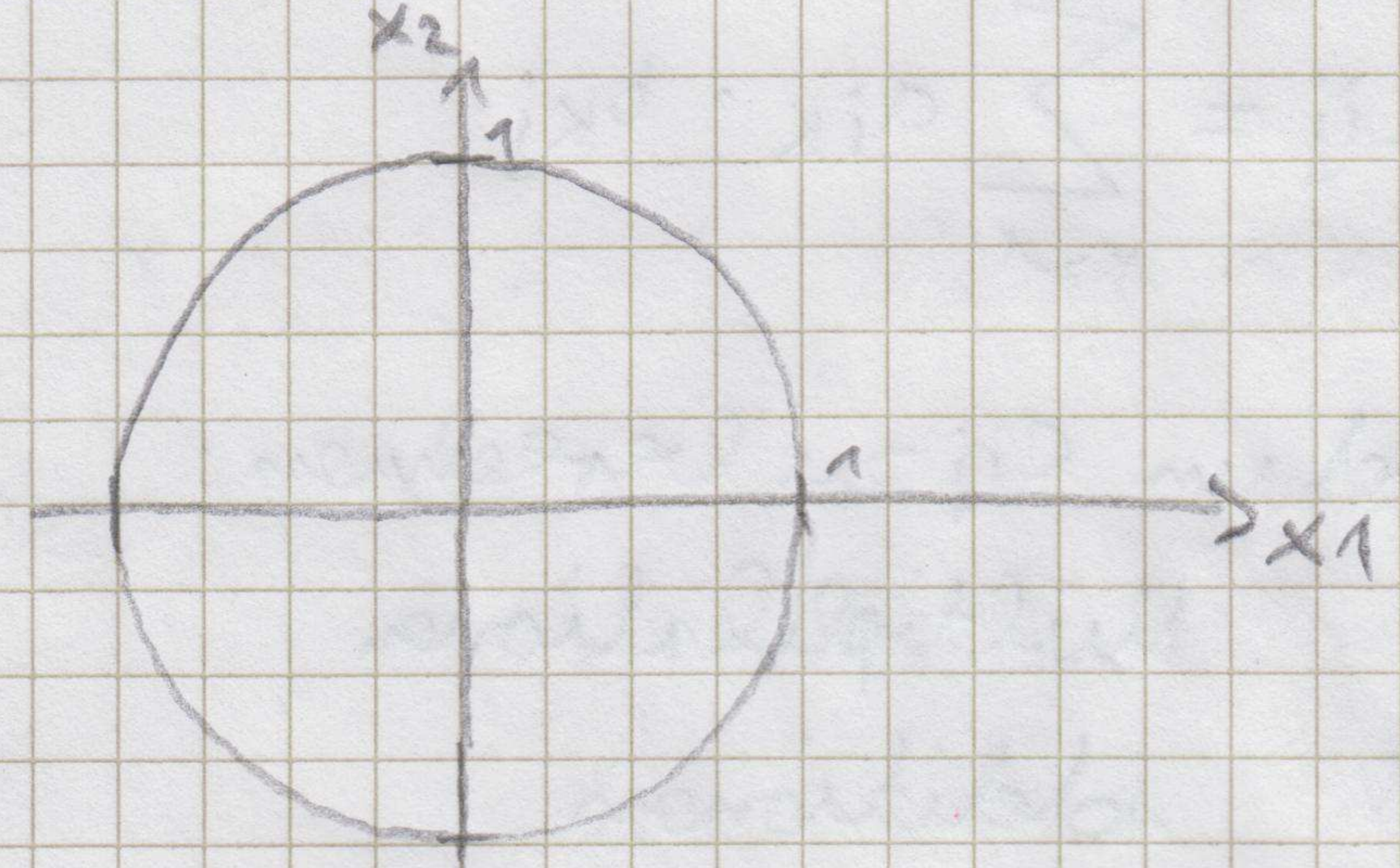
d) ∞-Norm (Maximums-Norm)

$$\|v\|_\infty = \max\{|1|, |2|, |3|\} = 3$$

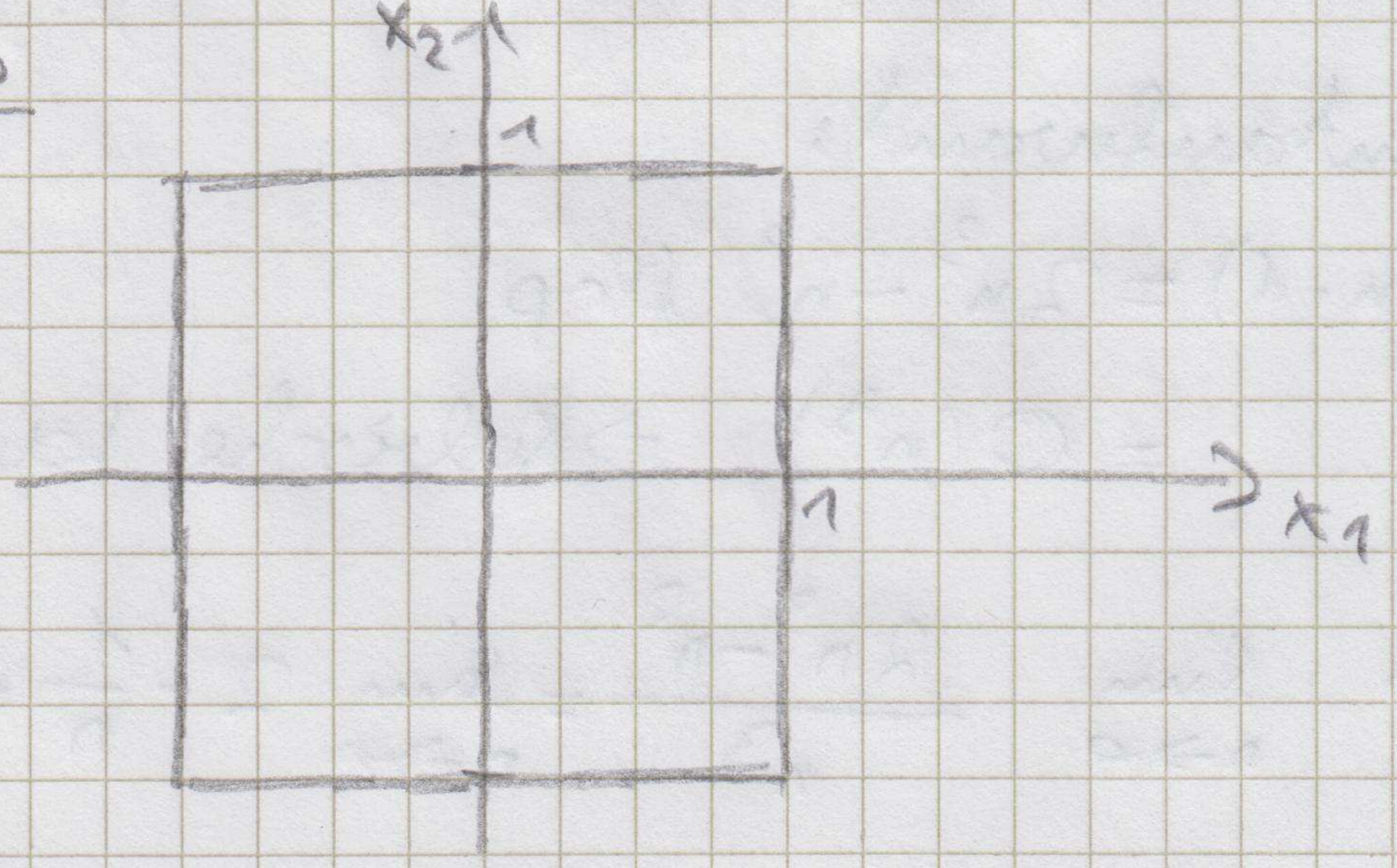
Einheitsball in der Ebene für verschiedene Normen

$$\{x \in \mathbb{R}^2 \mid \|x\|_p \leq 1\}$$

$\underline{p=2}$



$\underline{p=\infty}$



Bsp (direktes, inverses Problem, Parameteridentifikation)

a) direktes Problem:

$$\underset{\substack{\uparrow \\ \text{gesucht}}}{y} = f(x) = \underbrace{A \cdot x}_{\text{gegeben}} \quad \text{mit Drehmatrix } A = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}$$

b) inverses Problem:

LGS: $Ax = b$

$f: \mathbb{R}^n \to \mathbb{R}^n$

$$\underbrace{b = y = f(x) = A}_{\text{gegeben}} \cdot \underset{\substack{\uparrow \\ \text{gesucht}}}{x} \qquad \Rightarrow x = A^{-1} \cdot b$$

c) Parameter Identifikation

→ Computertomographie