Richard Huber
richard.huber@uni-graz.at

Aufgabe 7.1) [Median als Minimierungsproblem]
Es sei $(\Omega, \mathcal{A}, \mu)$ ein endlicher Maßraum, $f \in L^1_\mu(\Omega)$ und wir betrachten das Funktional

$$F\colon \mathbb{R} \to \mathbb{R}, \qquad F(a) = \|f - a\mathbf{1}\|_{L^1_\mu}, \tag{1}$$

wobei **1** die konstante Funktion mit Wert 1 bezeichnet.

a) Zeigen Sie, dass F differenzierbar in $a \in \mathbb{R}$, genau dann, wenn $\mu(\{f = a\}) = 0$, d.h. $f \neq a$ μ-fast überall. Bestimmen Sie darüber hinaus das Subdifferenzial $\partial F_{|a}$ für $a \in \mathbb{R}$.

b) Es bezeichne $L = \operatorname{argmin}_{a\in\mathbb{R}} F(a)$ die Menge der Minimierer. Folgern Sie, falls $\mu(\{f \leq a^*\}) = \mu(\{f \geq a^*\})$ so ist $a^* \in L$. Für $M = \{a \in L \mid \mu(\{f = a\}) \neq 0\}$, zeigen Sie dass jedes $a^* \in M$ ein extremaler Punkt von L ist, d.h. $a^* \in M$ ist $\max L$ oder $\min L$.

Bemerkung. *Wie das Beispiel zeigt, lässt sich der Median – wie Sie ihn vermutlich aus der Statistik Vorlesung kennen – als Minimierungsproblem schreiben, bzw. das Konzept des Medians zu komplexeren Maßen verallgemeinern. Dabei gilt die Bedingung dass "gleich viel Masse" oberhalb und unterhalb von a^* ist. Da das Problem nur 1-dimensional ist, ist es trivial zu zeigen, dass es Lösungen gibt, und man könnte eine Art "Bisektions-Algortihmus" nutzen um die Optimalitätsbedingungen zu lösen.*

Aufgabe 7.2) [Diskretes L^2-TV-Entrauschen]
Es sei $N \in \mathbb{N}$, $\Omega = \{1, \dots, N\} \times \{1, \dots, N\}$, $\lambda > 0$, $U^0 \in l^2(\Omega)$ und wir betrachten das (diskrete) Minimierungsproblem

$$\min_{U\in l^2(\Omega)} J(U) \qquad \text{mit} \qquad J(U) := \frac{\|U - U^0\|^2_{l^2(\Omega)}}{2} + \lambda \|\nabla_N U\|_{l^1(\Omega,\mathbb{R}^2)} \text{ für } U \in l^2(\Omega). \tag{2}$$

Hierbei bezeichnet für $v \in l^2(\Omega, \mathbb{R}^2)$, $\|v\|_{l^1(\Omega,\mathbb{R}^2)} = \sum_{ij} |v_{ij}|$ mit $|v_{ij}| = \sqrt{v_{ij1}^2 + v_{ij2}^2}$ (also l^1 mit punktweise euklidische Norm) und bezeichne $\|v\|_{l^\infty((\Omega,\mathbb{R}^2)} = \max_{ij} |v_{ij}|$ die duale Norm zu $\|\cdot\|_{l^1(\Omega,\mathbb{R}^2)}$. Es bezeichnet $\langle \cdot, \cdot \rangle$ das standard Skalarprodukt in l^2, was aber auch der Dualpaarung von l^∞ mit l^1 entspricht.

a) Zeigen Sie, dass (2) äquivalent zu

$$\min_{U\in l^2(\Omega)} \sup_{P\in l^2(\Omega,\mathbb{R}^2),\ \|P\|_\infty \leq \lambda} \frac{\|U - U^0\|^2_{l^2(\Omega)}}{2} - \langle \operatorname{div}_N P, U \rangle \tag{3}$$

ist, wobei $-\operatorname{div}_N = (\nabla_N)^*$ eine Approximation der Divergenz ist. Zeigen Sie unter der Annahme, dass in (3) die Reihenfolge des Minimums und Supremums vertauscht werden darf ohne den Wert zu ändern, dass (3) äquivalent zu

$$\min_{P\in l^2(\Omega,\mathbb{R}^2),\ \|P\|_\infty \leq \lambda} \frac{\|\operatorname{div}_N P + U^0\|^2_{l^2(\Omega)}}{2} \tag{4}$$

ist, und bestimmen Sie, wie optimales P^* in (4) mit optimalem U^* in (2) zusammenhängt.

b) Bestimmen Sie den Normalen-Kegel der Menge $\{P \in l^2(\Omega,\mathbb{R}^2) \mid \|P\|_\infty \leq \lambda\}$. Folgern Sie, dass P^* das Funktional in (4) minimiert, genau dann wenn es ein $\mu^* \in l^2(\Omega)$ gibt, sodass

$$\begin{cases} -\nabla_N\big(\mathrm{div}_N[P^*] + U^0\big) + \mu^* P^* = 0, \\ \mu^* \geq 0, \qquad \|P^*\|_\infty \leq \lambda, \\ \mu^*_{ij} = 0 \qquad \text{falls } |P^*_{ij}| < \lambda. \end{cases} \tag{5}$$

Dabei bezeichnet $\mu^* P^*$ komponentenweise Multiplikation und die Bedingung an μ^*_{ij} für alle $i,j \in \{1,\dots,N\}$.

Bemerkung. *Das Problem* (2) *ist nicht differenzierbar, weshalb Standard-Minimierungsalgorithmen (wie "Gradient-Descent" oder "Newton-Iteration") nicht anwendbar sind. Auch die direkten Optimalitätsbedingungen besitzen eine "unschöne" Struktur, und deren Lösung ist nicht leicht zu berechnen. Daher reformuliert man das Problem zum "dualen" Problem* (4)*, deren Optimalitätsbedingungen* (5) *eine bessere Struktur besitzen. Insbesondere sieht man, dass* μ^* *ein "Lagrange-Multiplikator" der Nebenbedingung ist.*

Programmier-Aufgabe 7.3) [Implementierung von L^2-TV-Entrauschen]
Wir definieren den Projektionsoperator

$$\mathrm{proj}_\lambda \colon \mathbb{R}^2 \to \mathbb{R}^2 \quad \text{mit} \qquad \mathrm{proj}_\lambda(q) = \begin{cases} \lambda \frac{q}{|q|} & \text{falls } |q| > \lambda, \\ q & \text{sonst}, \end{cases} \tag{6}$$

und verstehen $\mathrm{proj}_\lambda(P)$ für $P \in l^2(\Omega,\mathbb{R}^2)$ stillschweigend als komponentenweise Anwendung von proj_λ. Man kann zeigen, dass für beliebiges $\tau > 0$ die Optimalitätsbedingungen (5) äquivalent zu

$$P^* = \mathrm{proj}_\lambda\big(P^* + \tau\nabla_N(\mathrm{div}[P^*] + U^0)\big) \tag{7}$$

ist. Daraus lässt sich die Fixpunktiteration

$$\begin{cases} P_0 \in \mathbb{R}^{N\times M\times 2} \\ U_{n+1} = \mathrm{div}[P_n] + U^0 \\ P_{n+1} = \mathrm{proj}_\lambda\big(P_n + \tau\nabla_N(U_{n+1})\big) \end{cases} \tag{8}$$

ableiten. Für $\tau \leq \frac{1}{4}$ gilt $\mathrm{div}[P_n] + U^0 \to U^*$ mit U^* Lösung von (2).

Implementieren Sie die Iteration aus (8) mit $P_0 = 0$ und testen Sie sie anhand von U^0 entsprechend des Bildes Link mit $\lambda = \{0.1, 0.3, 1, 3\}$ und $\tau = \frac{1}{4}$ mit 100 Iterationen.

Bemerkung. *Wie in Beispiel 7.2 angedeutet, hat das duale Problem eine analytisch vorteilhaftere Struktur, da man Nicht-Differenzierbarkeit gegen Nebenbedingungen getauscht hat. Da Optimierung unter Nebenbedingungen aber ein überaus relevantes Feld ist, gibt es zahlreiche Algorithmen die auf die Lösung solcher Problem anwendbar sind, insbesondere entspricht obiger Algorithmus dem "Projected Gradient Descent".*

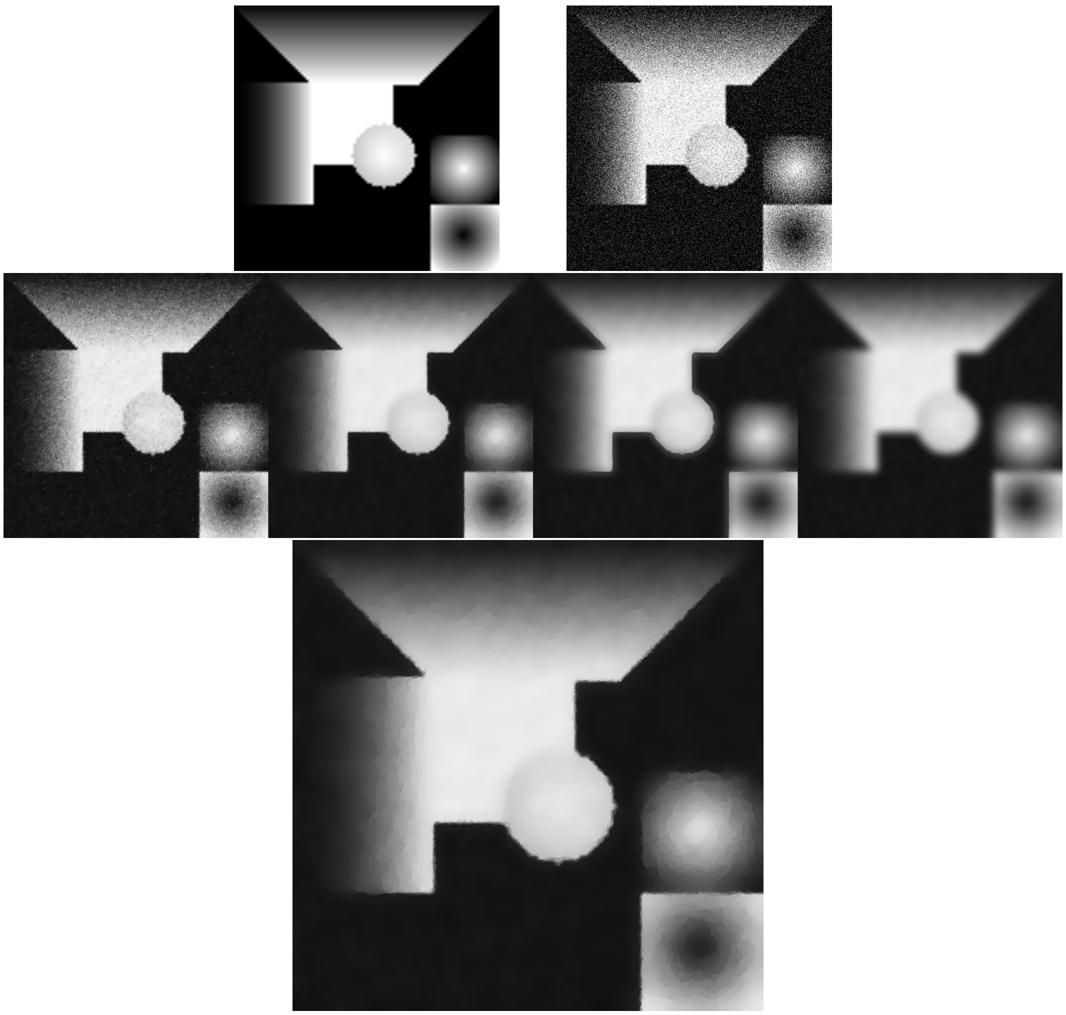

Abbildung 1: Erste Zeile zeigt Original-Bild, sowie Bild mit Gauss-Rauschen mit Standardabweichung 0.2. Zweite Zeile zeigt Rekonstruktionen für $\lambda = \{0.1, 0.3, 1, 3\}$ und 3. Zeile nochmal in gross den Fall $\lambda = 0.3$. Es gilt analog zur Bemerkung in Beispiel 5.3, dass die Wahl des Parameters λ Einfluss nimmt. Auffallend ist aber die Tatsache, dass die harten Übergänge (Kanten) sehr gut rekonstruiert werden und nicht verschwimmen (wie es bei L^2-H^1 Entrauschen der Fall war). Dies liegt daran, dass dieser Ansatz stückweise konstante Lösungen generiert. Dafür zahlt man aber den Preis, dass auch nicht-konstante Regionen konstant approximiert werden, was zu systematischen Artefakten (staircasing) führen kann. Dies sieht man beispielsweise in der Rekonstruktion mit $\lambda = 0.3$ an den Rampen links (quadratisch abfallend) und oben (linear abfallend), und in den Kugeln rechts unten (quadratisch in 2d).