Übungen zur Vorlesung
Optimierung bei partiellen Differentialgleichungen

G10. Semiglatte Newton-Verfahren

a) Bestimmen Sie für die Abbildung $G: \quad (x, y) \in \mathbb{R}^2 \mapsto \max(x, y) \in \mathbb{R}$ das Clarkesche Differential $\partial^{cl} G(x, y)$. Weisen Sie die $\partial^{cl} G$-Semiglattheit von G nach.

b) Betrachte das Optimierungsproblem

$$\min \ \hat{f}(u) \ \text{ s. t. } \ u \geq a$$

mit $u \in U := \mathbb{R}^m$, $\hat{f}: \mathbb{R}^m \to \mathbb{R}$ zweimal stetig differenzierbar, $a \in \mathbb{R}^m$. Die notwendigen Optimalitätsbedingungen lauten bekanntlich

$$\bar{u} \geq a, \quad (\nabla \hat{f}(\bar{u}))^T (u - \bar{u}) \geq 0 \ \ \forall\, u \geq a$$

oder äquivalent

$$\begin{pmatrix} \nabla \hat{f}(\bar{u}) - \bar{z} \\ \bar{u} - P(\bar{u} - \gamma \bar{z}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{O}$$

wobei $\gamma > 0$ beliebig ist und $P(u) = \max(u, a)$ die Projektion auf $\{u : u \geq a\}$ bezeichnet.

Zeigen Sie mit a): Der Operator links in (O) ist semiglatt. Ist $\nabla^2 \hat{f}(\bar{u})$ positiv definit, dann konvergiert das semiglatte Newton-Verfahren lokal q-superlinear gegen $(\bar{u}, \bar{z})$.

G11. Projizierte Gradientenschritte sind Abstiegsrichtungen

Betrachte das Problem

$$\min_{w \in W} f(w) \quad w \in \mathcal{S} \tag{P}$$

mit einem Hilbert-Raum W, $\mathcal{S} \subset W$ nichtleer, abgeschlossen und konvex und $f: W \to \mathbb{R}$ stetig F-differenzierbar. Sei $\nabla f(w) \in W$ die Riesz-Darstellung von $f'(w) \in W^*$.

Sei $P: W \to \mathcal{S}$ die Projektion auf $\mathcal{S}$. Um (P) zu lösen, betrachten wir das *projizierte Gradientenverfahren*:

0. Wähle $\delta \in]0, 1/2[$ und einen Startpunkt $w_0 \in \mathcal{S}$.

   Für $k = 0, 1, \ldots$:

1. Falls $\|w_k - P(w_k - \nabla f(w_k))\|_W = 0$: STOP mit stationärem Punkt w_k.
2. Bestimme das maximale $\sigma_k \in \{1, 2^{-1}, 2^{-2}, \ldots\}$, so dass $w_k(\sigma_k) := P(w_k - \sigma_k \nabla f(w_k))$ die Armijo-Bedingung erfüllt

   $$f(w_k) - f(w_k(\sigma_k)) \geq -\delta\, (\nabla f(w_k), w_k(\sigma_k) - w_k)_W.$$

3. Setze $w_{k+1} := w_k(\sigma_k)$.

Der wesentliche Punkt, warum dieses Verfahren funktioniert, ist die Tatsache, dass der projizierte Gradientenschritt $w_k(\sigma_k) - w_k = P(w_k - \sigma_k \nabla f(w_k)) - w_k$ eine *Abstiegsrichtung* ist, also gilt

$$(\nabla f(w_k), w_k(\sigma_k) - w_k)_W < 0 \quad \text{solange } w_k(\sigma_k) - w_k \neq 0,\, \sigma_k > 0.$$

Zeigen Sie mit Hilfe von Lemma 5.1.2:

a) Für jedes $w \in \mathcal{S}$ gilt

$$(\nabla f(w), w(\sigma) - w)_W < 0 \quad \text{solange } w(\sigma) - w \neq 0,\, \sigma > 0.$$

b) Es gilt sogar

$$(\nabla f(w), w(\sigma) - w)_W \leq \frac{-1}{\sigma} \|w(t) - w\|_W^2 \leq -\|w(1) - w\|_W \|w(\sigma) - w\|_W \quad \forall\, \sigma \in]0, 1].$$

c*) Warum ist also die Armijo-Bedingung in Schritt 2 für $\sigma_k = 2^{-j_k}$ klein genug erfüllt?