141. Wir betrachten nochmals das Gauss-Newton Verfahren aus Aufgabe 139 vom letzten Übungsblatt.

(a) Das Verfahren erzeuge eine gegen x^* konvergente Folge $\{x^k\}$. Weiters sei $F(x^*) = 0$ und $J_F(x^*)$ habe vollen Spaltenrang. Interpretieren Sie das Gauss-Newton-Verfahren als Newton-artiges Verfahren und zeigen Sie, dass es unter den angegeben Bedingungen q-superlinear gegen x^* konvergiert.

(b) Betrachten Sie nun den Fall $F : \mathbb{R} \to \mathbb{R}^2$ mit

$$F(x) = \begin{pmatrix} x \\ x^2/4 + 1 \end{pmatrix}.$$

Bestimmen Sie die globale Lösung x^* von (KQ), zeigen Sie dass das Gauss-Newton-Verfahren für $0 < |x^0| < 2$ gegen x^* konvergiert und untersuchen Sie das Konvergenzverhalten? Wieso ist Teil (a) nicht anwendbar?

142. Beweisen Sie die beiden folgenden Hilfsaussagen, die in der Herleitung von Quasi-Newton-Formeln in der Vorlesung verwendet wurden:

(a) Für alle $w \in \mathbb{R}^n$ gilt $||w|| = \max_{||x||=1} |w^t x|$.

(b) Für alle $v, w \in \mathbb{R}^n$ gilt $||vw^t|| = ||v||\,||w||$.

143. Seien $f : \mathbb{R}^n \mapsto \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, $x^* \in \mathbb{R}^n$, $\{H^{(k)}\}$ eine Folge von symmetrischen Matrizen über $\mathbb{R}$ und $\{x^{(k)}\}$ eine Folge die gegen x^* superlinear konvergiert. Zeigen Sie dass, die folgenden beiden Aussagen zueinander äquivalent sind:

(a) Es gilt $||(\nabla^2 f(x^{(k)}) - H^{(k)})(x^{(k+1)} - x^{(k)})|| = o(||x^{(k+1)} - x^{(k)})||$.

(b) Es gilt $||\nabla f(x^{(k+1)}) - \nabla f(x^{(k)}) - H^{(k)}(x^{(k+1)} - x^{(k)})|| = o(||x^{(k+1)} - x^{(k)})||$.

144. Gegegeben sei das quadratische Optimierungproblem min $\frac{1}{2}x^t Q x + c^t x + \gamma$ wobei $c \in \mathbb{R}^n$ und Q eine symmetrische, positiv definite $n \times n$ Matrix über $\mathbb{R}$ ist.

Betrachten Sie das lokale Quasi Newton-Verfahren mit SR1 (symmetrischer Rang 1) Update und wenden Sie es auf f an.

(0) Wähle $x^{(0)} \in \mathbb{R}^n$, und $H^{(0)}$ symmetrisch und invertierbar. $k := 0$

(1) Falls $\nabla f(x^{(k)}) = 0$, STOP.

(2) Berechne $d^{(k)}$ durch Lösen von

$$H^{(k)} d^{(k)} = -\nabla f(x^{(k)})$$

(3) Setze $x^{(k+1)} = x^{(k)} + d^{(k)}$, $s^{(k)} = d^{(k)}$, $y^{(k)} = \nabla f(x^{(k+1)}) - \nabla f(x^{(k)})$

(4) Setze

$$H^{(k+1)} = H^{(k)} + \frac{(y^{(k)} - H^{(k)}s^{(k)})(y^{(k)} - H^{(k)}s^{(k)})^t}{(y^{(k)} - H^{(k)}s^{(k)})^t s^{(k)}}$$

(5) Setze $k := k + 1$ und gehe zu (1).

Wir nehmen an, dass $(y^{(k)} - H^{(k)}s^{(k)})^t s^{(k)} \neq 0$ für alle k gilt. Zeigen Sie, dass dann folgendes gilt

(a) $H^{(k)}s^{(j)} = y^{(j)}$ für alle $j = 0, \ldots, k-1$.

(b) Das Verfahren findet spätestens im n-ten Schritt das Minimum.

(c) Wenn n Schritte ausgeführt werden und die zugehörigen n Suchrichtungen linear unabhängig sind, dann gilt $H^{(n)} = Q$ (d.h. $H^{(n)}$ approximiert die Hessematrix exakt).

(d) Aus $H^{(k)}$ positiv definit und $(s^{(k)})^t y^{(k)} > 0$ folgt nicht, dass $H^{(k+1)}$ positiv definit ist.

145. Gegeben sei das quadratische Optimierungsproblem min $\frac{1}{2}x^tQx + c^tx\gamma$ mit $n = 3$,

$$Q = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad c = \begin{pmatrix} 8 \\ 9 \\ 8 \end{pmatrix}$$

Starten Sie mit $x^{(0)} = (0,0,0)^t$ und $H^{(o)} = I$. Wenden Sie das in Aufgabe 144 beschriebene Verfahren an.

146. Gegeben seien die $n \times m$ Matrizen U, V über $\mathbb{R}$ sowie eine $n \times n$ Matrix A über $\mathbb{R}$ und eine $m \times m$ Matrix über $\mathbb{R}$. A und S seien invertierbar.

   (a) Zeigen Sie dass die Matrix M mit $M := A + USV^t$ genau dann invertierbar ist, wenn $W := S^{-1} + V^tA^{-1}U$ invertierbar ist und dass im Fall der Existenz von M^{-1} folgendes gilt:

$$M^{-1} = A^{-1} - A^{-1}UW^{-1}V^tA^{-1}.$$

   (b) Übertragen Sie den Fall aus 146a, um die in der Vorlesung erwähnte Resultat für den Fall einer Rang-1-Modifikation $M := A + uv^t$ mit $u, v \in \mathbb{R}^n$ zu zeigen: M ist invertierbar genau dann wenn $1 + v^tA^{-1}u \neq 0$ gilt und dann gilt

$$M^{-1} = \left(I - \frac{A^{-1}uv^t}{1 + v^tA^{-1}u}\right)A^{-1}.$$

   (c) Verwenden Sie die Formel aus Teil 146b, um die zur Updateformel aus Aufgabe 144 inverse symmetrische Rang 1 Updateformel

$$(H^{(k+1)})^{-1} = B^{(k+1)} = B^{(k)} + \frac{(s^{(k)} - B^{(k)}y^{(k)})(s^{(k)} - B^{(k)}y^{(k)})^t}{(s^{(k)} - B^{(k)}y^{(k)})^t y^{(k)}}$$

   herzuleiten.

147. Sei H eine symmetrische $n \times n$ Matrix über $\mathbb{R}$ und seien $u \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}$.

   Der Update $H^+ = H + \gamma uu^t$ wird symmetrischer Rang-1-Update genannt. Zeigen Sie, dass sich als eindeutige Lösung der Quasi-Newtongleichung $H^+s = y$ folgende Formel für den Rang-1-Update ergibt

$$H^+ = H + \frac{(y - Hs)(y - Hs)^t}{(y - Hs)^t s}$$

   sofern $(y - Hs)^t s \neq 0$ gilt.

148. Sei H eine symmetrische und positive definite $n \times n$ Matrix über $\mathbb{R}$ und seien $u, v \in \mathbb{R}^n$ und $\gamma, \delta \in \mathbb{R}$. Zeigen Sie, dass man die direkte BFGS-Formel als Rang-2-Update der Form $H^+ = H + \gamma u^tu + \delta v^tv$ interpretieren kann.

149. Wenden Sie das BFGS-Quasi Newtonverfahren auf das Beispiel aus Aufgabe 145 an. Wieviele Schritte werden benötigt? Wie ist dies zu erklären?

150. Rechnen Sie nach, dass $H^{(k+1)}s^{(k)} = y^{(k)}$ gilt für $H^{(k+1)}$ aus der (direkten) BFGS-Updateformel

$$H^{(k+1)} = H^{(k)} + \frac{y^{(k)}(y^{(k)})^t}{(y^{(k)})^t s^{(k)}} - \frac{H^{(k)}s^{(k)}(H^{(k)}s^{(k)})^t}{s^{(k)}H^{(k)}(s^{(k)})^t}.$$

151. Sei wiederum $H^{(k+1)}$ aus $H^{(k)}$ mittels der BFGS-Update Regel erzeugt.

(a) Zeigen Sie, dass $H^{(k+1)}$ positiv definit ist, wenn $(y^k)^t s^k > 0$.

(b) Zeigen Sie, dass wenn die Wolfe-Powell-Bedingung Schrittweitenstrategie mit $\rho < 1$ verwendet wird, und somit die gewählte Schrittweite t_k

$$|d^k \nabla f(x^k + t_k d^k))| \leq \rho |(d^k)^t \nabla f(x^k)|$$

erfüllt, dass dann $(y^k)^t s^k > 0$ gilt und somit die Matrix H_{k+1} positiv definit ist, wenn H_k positiv definit ist.