Albert-Ludwigs-Universität
Institut für Informatik
Prof. Dr. F. Kuhn, Oleksii Saukh 21.06.2013

Einführung in die Optimierung

Sommersemester 2013

Übungsblatt 1 – Musterlösung

Übung 1: Konvexität

Die Hesse'sche Matrix $H(\boldsymbol{x})$ einer quadratischen Funktion ist an allen Stellen $\boldsymbol{x} = (x, y, z)^\top$ identisch. Für die gegebene Funktion f gilt:

$$H(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial^2}{\partial x^2} f(x,y) & \frac{\partial^2}{\partial x \partial y} f(x,y) \\ \frac{\partial^2}{\partial x \partial y} f(x,y) & \frac{\partial^2}{\partial y^2} f(x,y) \end{pmatrix} = \beta \cdot \begin{pmatrix} 2 & -\alpha \\ -\alpha & 8 \end{pmatrix} =: H.$$

Die Matrix H is positiv semidefinit falls für alle Vektoren $\boldsymbol{u} = (u, v)^\top$ gilt, dass $\boldsymbol{u}^\top H \boldsymbol{u} \geq 0$. Das heisst, folgende Ungleichung muss für alle $u, v \in \mathbb{R}$ erfüllt sein:

$$\forall u, v \in \mathbb{R} : \ 2\beta(u^2 + 4v^2 - \alpha uv) \geq 0. \tag{1}$$

Schauen wir uns zuerst den Ausdruck $u^2 + 4v^2 - \alpha uv$ in der Klammer an. Falls genau eine der Variablen u und v 0 ist, ist der Ausdruck positiv. Der Ausdruck kann daher für alle α positiv werden und wir müssen untersuchen, für welche α er auch negativ werden kann. Für festes v beschreibt der Ausdruck eine quadratische Funktion in u, welche bei $u = \alpha v/2$ das globale Minimum hat. Für festes v ist das Minimum des Ausdrucks daher

$$\frac{\alpha^2 v^2}{4} + 4v^2 - \frac{\alpha^2 v^2}{2} = v^2 \left(4 - \frac{\alpha^2}{4}\right).$$

Dieser Wert ist negativ genau dann, wenn $\alpha^2 > 16$. Falls $-4 \leq \alpha \leq 4$ kann der Ausdruck $u^2 + 4v^2 - \alpha uv$ daher nicht negativ werden. Falls $|\alpha| > 4$ kann der Ausdruck sowohl positive, als auch negative Werte annehmen. Gleichung (1) ist deshalb erfüllt und f damit konvex genau dann, wenn $\beta \geq 0$ und $-4 \leq \alpha \leq 4$. Für die Funktion $-f$ erhalten wir als Hesse'sche Matrix $-H$ und damit muss der Ausdruck in Gleichung (1) kleiner oder gleich 0 sein, damit $-f$ konvex und damit f konkav ist. Dies ist erfüllt genau dann, wenn $\beta \leq 0$ und $-4 \leq \alpha \leq 4$ ist.

Übung 2: Gradientenverfahren

In Matrix-Schreibweise können wir die Funktion $f(x, y, z) = x^2 + \frac{1}{2}y^2 + 2z^2 - xy + yz - \frac{2}{3}x - \frac{2}{3}y - 6z$ folgendermassen aufschreiben:

$$f(x, y, z) = \frac{1}{2} \cdot \boldsymbol{x}^\top A \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x}, \quad \text{wobei } A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 4 \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 2/3 \\ 2/3 \\ 6 \end{pmatrix}.$$

Als Gradient bekommen wir damit $\nabla f(\boldsymbol{x}) = A\boldsymbol{x} - \boldsymbol{b}$, und für die Abstiegsrichtung $\boldsymbol{d_0}$ and der Stelle $\boldsymbol{x_0} = (x_0, y_0, z_0)^\top = (4/3, 1, 1)^\top$ erhalten wir

$$\boldsymbol{d_0} = -\nabla f(\boldsymbol{x_0}) = \boldsymbol{b} - A\boldsymbol{x_0} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Um die optimale Schrittweite τ_0 zu finden, müssen wir daher das eindimensionale Minimierungsproblem

$$\arg\min_{\tau_0 \in \mathbb{R}} f(\boldsymbol{x_0} + \tau_0 \cdot \boldsymbol{d_0})$$

lösen. Wir erhalten ein eindimensionales quadratisches Minimierungsproblem, welches wir lösen können, indem wir nach τ_0 ableiten und den erhaltenen Ausdruck 0 setzen:

$$0 \stackrel{!}{=} \frac{\partial}{\partial \tau_0} f(\boldsymbol{x_0} + \tau_0 \boldsymbol{d_0}) = \boldsymbol{d_0}^\top (A(\boldsymbol{x_0} + \tau_0 \boldsymbol{d_0}) - \boldsymbol{b}) = \tau_0 \boldsymbol{d_0}^\top A \boldsymbol{d_0} + \boldsymbol{d_0}^\top (A\boldsymbol{x_0} - \boldsymbol{b}).$$

Für die optimale Schrittweite τ_0 erhalten wir daher

$$\tau_0 = \frac{-\boldsymbol{d_0}^\top (A\boldsymbol{x_0} - \boldsymbol{b})}{\boldsymbol{d_0}^\top A \boldsymbol{d_0}} = \frac{-\boldsymbol{d_0}^\top \nabla f(\boldsymbol{x_0})}{\boldsymbol{d_0}^\top A \boldsymbol{d_0}} = \frac{\boldsymbol{d_0}^\top \boldsymbol{d_0}}{\boldsymbol{d_0}^\top A \boldsymbol{d_0}} = \frac{2}{6} = \frac{1}{3}.$$

Als Startwert für den nächsten Iterationsschritt erhalten wir dann

$$\boldsymbol{x_1} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \boldsymbol{x_0} + \tau_0 \boldsymbol{d_0} = \begin{pmatrix} 4/3 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3} \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 4/3 \end{pmatrix}.$$

Für den zweiten Iterationsschritt erhalten wir analog

$$\boldsymbol{d_1} = \boldsymbol{b} - A\boldsymbol{x_1} = \begin{pmatrix} -1/3 \\ -2/3 \\ -1/3 \end{pmatrix}, \quad \text{sowie} \quad \tau_1 = \frac{\boldsymbol{d_1}^\top \boldsymbol{d_1}}{\boldsymbol{d_1}^\top A \boldsymbol{d_1}} = \frac{3}{5}$$

und damit

$$\boldsymbol{x_2} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \boldsymbol{x_1} + \tau_1 \boldsymbol{d_1} = \begin{pmatrix} 1 \\ 1 \\ 4/3 \end{pmatrix} + \frac{3}{5} \cdot \begin{pmatrix} -1/3 \\ -2/3 \\ -1/3 \end{pmatrix} = \begin{pmatrix} 4/5 \\ 3/5 \\ 17/15 \end{pmatrix}.$$

Die Funktionswerte an den Stellen $\boldsymbol{x_0}$, $\boldsymbol{x_1}$ und $\boldsymbol{x_2}$ sind

$$f(\boldsymbol{x_0}) = -3.6111, \quad f(\boldsymbol{x_1}) = -3.9444 \quad \text{und} \quad f(\boldsymbol{x_2}) = -4.1444.$$