Solutions

Question 1  [3 marks] Indicate, for each of the following statements, whether it is true or false. Write “True” or “False” next to each item. No explanation is necessary.

- If a function \( f \in C^1(\mathbb{R}^n) \) has a unique global minimizer, then it is strictly convex. **False**, not necessarily, convexity is not even needed. Take \(-\sin(x)/x\).

- If \( f \in C^2(\mathbb{R}^n) \) and \( x^* \) is a strict local minimizer for \( f \) then \( \nabla f(x^*) = 0 \) and \( \nabla^2 f(x^*) \) is positive definite. **False**, consider \( f(x) = x^4 \) at \( x^* = 0 \).

- Let \( f \in C^0(\mathbb{R}^n) \) and let \( L = \{x \in \mathbb{R}^n : f(x) \leq 2\} \). If \( L \neq \emptyset \) and \( L \subseteq B_1(0) \), then \( f \) has a global minimizer. **True**, because this is stating that \( f \) has a nonempty bounded level set.

- Let \( f : \mathbb{R}^n \to \mathbb{R} \), and let \( p \in \mathbb{R}^n \) be a descent direction at \( x \in \mathbb{R}^n \). If \( f(x + \alpha p) > f(x) + \sigma \alpha \nabla f(x)^T p \), then \( \sigma \) satisfies the curvature condition (where \( 0 < \sigma < \frac{1}{2} \)) for line search. **True**, this is the definition.

- If the trust region method converges to a point \( x^* \), then \( ||x^* - x^k||_2 \leq \delta^k \) at any iteration \( k \) where \( x^k \) is the current point and \( \delta^k \) is the current trust region radius. **False**, \( ||x^* - x^0||_2 \leq \delta \) is not necessary for convergence. \( \delta \) does not play any role for convergence in general, only for quadratic convergence.

Question 2  [3 marks] Let \( f : \mathbb{R}^n \to \mathbb{R} \) be defined by \( f(x) = x^T (A + \beta I)x + b^Tx \), where \( b \in \mathbb{R}^n \), \( \beta \in \mathbb{R} \), and \( A \in \mathbb{R}^{n \times n} \) is a symmetric matrix that is not positive semidefinite (i.e. \( A \) has at least one negative eigenvalue). Prove that if \( x^* \) is a global minimizer for \( f \) and \( ||x^*||_2 = 1 \), then \( x^* \) is an optimal solution to

\[
\min \quad x^T Ax + b^T x \\
\text{s.t.} \quad ||x||_2 \leq 1.
\]

**Solution:** First, \( \min\{x^T Ax + b^T x : ||x||_2 \leq 1\} \) cannot have a global minimizer \( \bar{x} \) such that \( ||\bar{x}||_2 < 1 \), because then \( \bar{x} \) would be a local minimizer for \( x^T Ax + b^T x \), which is impossible since \( A \) is not positive semidefinite. Thus any global minimizer \( \bar{x} \) to \( \min\{x^T Ax + b^T x : ||x||_2 \leq 1\} \) must satisfy \( ||\bar{x}||_2 = 1 \).
Then, $x^*$ is an optimal solution to the each of the following problems:

\[
\begin{align*}
\min & \left\{ x^T (A + \beta I)x + b^T x : x \in \mathbb{R}^n \right\} \\
\min & \left\{ x^T (A + \beta I)x + b^T x : ||x||_2 = 1 \right\} \\
\min & \left\{ x^T Ax + x^T \beta x + b^T x : ||x||_2 = 1 \right\} \\
\min & \left\{ x^T Ax + \beta + b^T x : ||x||_2 = 1 \right\} \\
\min & \left\{ x^T Ax + b^T x : ||x||_2 = 1 \right\} \\
\min & \left\{ x^T Ax + b^T x : ||x||_2 \leq 1 \right\}
\end{align*}
\]

where we (1) restated the hypothesis, (2) restricted the feasible region to a subset that contains $x^*$, (3) distributed $(A + \beta I)$, (4) observed that $x^T x = ||x||_2^2 = 1$, (5) observed that $\beta$ is a constant, and (6) used $||\bar{x}||_2 = 1$ as stated above.

**Question 3** [4 marks] Consider a function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 - 6x^2 + 9x$.

(i) Determine all values of $x^k$ such that $x^{k+1}$ is well defined with Newton’s method, and find a local minimizer $x^*$ for $f$.

(ii) Let $x^*$ be the local minimizer found above. Prove that if $x^0 \in B_r(x^*)$ with $r = \frac{1}{2}$, then Newton’s method will converge quadratically to $x^*$.

**Solution:**

(i) We have

\[
\begin{align*}
f(x) &= x^3 - 6x^2 + 9x \\
f'(x) &= 3x^2 - 12x + 9 \\
f''(x) &= 6x - 12
\end{align*}
\]

Newton’s method is defined if $\nabla^2 f$ is positive definite, i.e. if $6x - 12 > 0$, which is for all $x^k > 2$. The critical points of $f$ are the roots of $f'$, i.e. $x = 1$ and $x = 3$. The Hessian is negative at $x = 1$, but its value is $\nabla^2 f(3) = 1$ at $x^* = 3$. Thus $x^* = 3$ is a (strict) local minimizer.

(ii) We have seen in class a theorem for the convergence of Newton’s method. We need to show that all its hypotheses hold.

(1) $\nabla^2 f$ is Lipschitz continuous over $B_r(x^*)$. We have that $||(6x - 12) - (6y - 12)|| \leq L(x - y)$ with $L = 6$. So $\nabla^2 f$ is Lipschitz continuous over $\mathbb{R}$.

(2) We checked in (i) that the second order sufficient conditions for optimality are satisfied at $x^* = 3$.

(3) $||\nabla^2 f(x)^{-1}|| \leq \frac{2}{\nabla^2 f(x^*)^{-1}}$. We have $\frac{1}{\nabla^2 f(x)} = \frac{1}{6(x-2)}$ and $\frac{1}{\nabla^2 f(x^*)} = \frac{1}{6}$. We must thus satisfy $x \geq \frac{5}{2}$, which is satisfied for all $r \leq \frac{1}{2}$.

(4) $r \leq \frac{1}{2L||\nabla^2 f(x^*)^{-1}||}$. This is true for

\[
r \leq \frac{1}{2 \cdot 6 \cdot |\frac{1}{6}|} = \frac{1}{2}.
\]