Exercises for week 9 discussion

The following exercises will be the topic of the discussion section.

1. *Perturbation lemma.* With the notation of page 16.3, show that if \( A \) is invertible and \( \|A^{-1}B\| < 1 \), then

\[
\|(A + B)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}B\|}{1 - \|A^{-1}B\|} \|A^{-1}\|.
\]

2. [Deuflhard] Consider a nonlinear equation \( f(x) = 0 \) where \( f : \mathbb{R}^n \to \mathbb{R}^n \) is differentiable. Suppose \( f(\hat{x}) \neq 0 \) and the Jacobian matrix \( f'(\hat{x}) \) of \( f \) at \( \hat{x} \) is nonsingular. Show that the Newton direction \( v = -f'(\hat{x})^{-1}f(\hat{x}) \) is a descent direction of the function \( g_A(x) = \|Af(x)\|_2^2 \), for any nonsingular matrix \( A \). In other words, show that

\[
\nabla g_A(\hat{x})^Tv < 0 \quad \text{for all nonsingular } A.
\] (1)

Are there other directions \( v \) (other than the Newton direction) with this property?

3. *Local convergence of Newton’s method.* The Kantorovich theorem in lecture 16 is called a semi-local convergence result, because it does not start from an assumption that a solution exists. If we make this assumption, local convergence is easier to prove. Let \( x^* \) be a solution at which \( f'(x^*) \) is invertible and define \( \alpha = \|f'(x^*)^{-1}\| \). Assume that \( f \) is differentiable in a neighborhood \( B = \{x \mid \|x - x^*\| < \rho\} \) and that \( f' \) is \( \beta \)-Lipschitz continuous on this set:

\[
\|f'(x) - f'(y)\| \leq \beta \|x - y\| \quad \text{for all } x, y \in B.
\]

(a) Suppose \( x_k \in B \) and \( \|x_k - x^*\| \leq 1/(2\alpha\beta) \). Use the perturbation result on page 16.3 to show that \( f'(x_k) \) is invertible and that

\[
\|f'(x_k)^{-1}\| \leq 2\alpha.
\]

(b) It can be shown that the Lipschitz continuity assumption implies that

\[
\|f(y) - f(x) - f'(x)(y - x)\| \leq \frac{\beta}{2}\|y - x\|^2 \quad \text{for all } x, y \in B.
\]

Use this fact to prove that

\[
\|x_{k+1} - x^*\| \leq \alpha\beta\|x_k - x^*\|^2.
\]

(c) Conclude that if \( x_0 \in B \) and \( \|x_0 - x^*\| \leq 1/(2\alpha\beta) \), then the iterates \( x_k \) in Newton’s method are well defined and converge to \( x^* \).