ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Prof. D. Kressner
M. Steinlechner

1 ► Higher-order derivatives

Calculating higher order derivatives of functions $f: \mathbb{R}^m \supset \Omega \to \mathbb{R}^d$ by calculating all partial derivatives is not always advisible. If $f \in C^{p+1}(\mathbb{R}^m, \mathbb{R}^d)$, then at a point $\mathbf{x} \in \Omega$ the *kth derivatives* are the *unique symmetric k-linear forms* $f^{(k)}(\mathbf{x})$ that satisfy

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \sum_{k=1}^{p} \frac{1}{k!} f^{(k)}(\mathbf{x})[\mathbf{h}, \dots, \mathbf{h}] + O(\|\mathbf{h}\|^{p+1}).$$

Note that in general the constant in $O(\|\mathbf{h}\|^{p+1})$ can be chosen independent of $\mathbf{x}$ in any bounded neighborhood of $\mathbf{x}$, but depends on the size of the neighborhood. To find $f^{(k)}(\mathbf{x})[\mathbf{h}, \dots, \mathbf{h}]$, one tries to expand $f(\mathbf{x}+\mathbf{h})$ and collect the terms of order $\|\mathbf{h}\|^k$.

Try it out by deriving $f'(\mathbf{x})[\mathbf{h}]$ and $f''(\mathbf{x})[\mathbf{h}, \mathbf{h}]$ for the following functions:

(a) $f: \mathbb{R}^n \to \mathbb{R},\ \mathbf{x} \mapsto \|\mathbf{x}\|_2^2$,

(b) $f: \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},\ A \mapsto A^3$.

(c) $f: \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n,\ \mathbf{x} \mapsto \frac{\mathbf{x}}{\|\mathbf{x}\|_2^2}$.

Hint. Use $(x+h)^{-1} = x^{-1} - x^{-2}h + \frac{1}{2}x^{-3}h^2 + O(|h|^3)$.

2 ► Validating derivatives calculated by hand

Since calculating derivatives is error-prone, it is good to have a numerical validation.

(a) Write a Matlab function `check_derivatives(f,f1,f2,x,N)` which calculates

$$f(\mathbf{x}+\mathbf{h}) - (f(\mathbf{x}) + f_1(\mathbf{x})[\mathbf{h}] + 1/2 f_2(\mathbf{x})[\mathbf{h}, \mathbf{h}])$$

for N different values of $\|\mathbf{h}\|$ tending appropriately to zero (you are free to choose whether the direction of $\mathbf{h}$ is fixed). The function should produce reasonable plots which allow to decide whether $f_1(\mathbf{x})$ and/or $f_2(\mathbf{x})$ are the first and second derivatives.

(b) We want to test the program with the function

$$f: \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},\ A \mapsto A^{-1}.$$

Candidates for $f'(A)[H]$ are

$$-A^{-2}H, \quad -A^{-1}HA^{-1}, \quad \text{or} \quad -HA^{-2},$$

while $f''(A)[H, H]$ could be

$$2HA^{-3}H, \quad \text{or} \quad 2A^{-1}HA^{-1}HA^{-1},$$

amongst other permutations. Which choices are correct?

Hint. For once use the Matlab function `inv` for the matrix inverse, which is usually not advisible.

(c)* Can you prove that the obtained expressions are indeed the derivatives?