Task 1.1: Gradient and Differentiability

Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is differentiable and define $F\colon \mathbb{R}^2 \to \mathbb{R}$ by setting

$$F(x,y) = f(x+2y) + f(7y-3x)$$

for all $x, y \in \mathbb{R}$. Is F differentiable? In that case, compute ∇F.

Set $g_1(x,y) = x+2y$ and $g_2(x,y) = 7y-3x$. As differentiability is a linear property, we know that g_1 and g_2 are differentiable on $\mathbb{R}^2$. Given f, g_1 and g_2 are differentiable, both $f \circ g_1$ and $f \circ g_2$ are differentiable, and thus through linearity, F is as well.

$$\begin{aligned}
\nabla F &= \frac{\partial F}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{bmatrix} \\
\frac{\partial F}{\partial x} &= \frac{\partial f \circ g_1}{\partial x} + \frac{\partial f \circ g_2}{\partial x} = \frac{\partial f}{\partial g_1}\frac{\partial g_1}{\partial x} + \frac{\partial f}{\partial g_2}\frac{\partial g_2}{\partial x} = f' \circ g_1 - 3f' \circ g_2 \\
\frac{\partial F}{\partial y} &= \frac{\partial f \circ g_1}{\partial y} + \frac{\partial f \circ g_2}{\partial y} = \frac{\partial f}{\partial g_1}\frac{\partial g_1}{\partial y} + \frac{\partial f}{\partial g_2}\frac{\partial g_2}{\partial y} = 2f' \circ g_1 + 7f' \circ g_2 \\
\Rightarrow \nabla F &= \begin{bmatrix} f' \circ g_1 - 3f' \circ g_2 \\ 2f' \circ g_1 + 7f' \circ g_2 \end{bmatrix} \Rightarrow \nabla F(\boldsymbol{x}) = \begin{bmatrix} f'(g_1(\boldsymbol{x})) - 3f'(g_2(\boldsymbol{x})) \\ 2f'(g_1(\boldsymbol{x})) + 7f'(g_2(\boldsymbol{x})) \end{bmatrix}
\end{aligned}$$

Task 1.2: Absolute Value

Define $F\colon \mathbb{R}^2 \to \mathbb{R}$ by setting $F(x,y) = \sqrt{|xy|}$. Is F differentiable in $\mathbf{0}$? Justify your answer.

Let $f_1(x) = \sqrt{x}$, $f_2(x) = |x|$ and $f_3(x,y) = xy$, then $F = f_1 \circ f_2 \circ f_3$. Given f_2 is not differentiable at $x = 0$, F is not differentiable at $\mathbf{0}$.

Recall on the ***Banach-Fixed-Point*** theorem. Let (X,d) be a non-empty complete metric space and $F\colon X \to X$. We say F is a *contraction* if there exists a constant $C \in [0,1)$ such that

$$d(F(z_1), F(z_2)) \leq C d(z_1, z_2)$$

for all $z_1, z_2 \in X$. We say that $\overline{z}$ ist a *fixed point* for F if $F(\overline{z}) = \overline{z}$. The Banach-Fixed-Point theorem states that if F is a contraction, then F admits a unique fixed point. Recall that $\mathbb{R}^n$ is a complete metric space with the euclidean distance.

Task 1.3: Banach-Fixed-Point Theorem

Let (X,d) be a non-empty complete metric space. Prove the Banach-Fixed-Point theorem stated above.

Let $x_0 \in X$ be arbitrary and F a contraction on X with constant $C \in [0,1)$. We define a sequence $(x_n)_{n\in\mathbb{N}}$ inductively where $x_{n+1} = F(x_n)$. We prove $(x_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence. Since X is complete under d, if $(x_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence, it converges:

$$d(x_l, x_{l+1}) = d(g(x_{l-1}), g(x_l)) \leq C d(x_{l-1}, x_l) \leq C^2 d(x_{l-2}, x_{l-1}) \leq \cdots \leq C^l d(x_0, x_1)$$

Therefore for some $l \in \mathbb{N}$ with $l \geq 1$, we know that $d(x_l, x_{l+1}) \leq C^l d(x_0, x_1)$, i.e. the distance between subsequent entries becomes 0 as $n \to \infty$. We now prove that $(x_n)_{n\in\mathbb{N}}$ is in fact a Cauchy-sequence. Let $n, m \in \mathbb{N}$ be arbitrary indices:

$$d(x_{n+m}, x_n) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+m-1}, x_{n+m})$$

$$\begin{aligned}
&\leq C^n d(x_0,x_1) + C^{n+1} d(x_0,x_1) + \cdots + C^{n+m-1} d(x_0,x_1) \\
&= C^n d(x_0,x_1) \sum_{j=0}^{m-1} C^j \leq C^n d(x_0,x_1) \sum_{j=0}^{\infty} C^j = d(x_0,x_1) \frac{C^n}{1-C}
\end{aligned}$$

Given $d(x_n, x_{n+m}) \leq d(x_0,x_1)\frac{C^n}{1-C} < \varepsilon$ we get

$$C^n < \frac{\varepsilon(1-C)}{d(x_0,x_1)+1}$$

As $C < 1$, we can choose n arbitrarily large, hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence, as the distance of any two entries for large n diminishes. Given F is Lipschitz-continuous with Lipschitz-constant C, we can apply the sequence-criterion:

$$\lim_{n\to\infty} F(x_n) = F(\overline{x})$$

Let $\overline{x}, \overline{y}$ be fixed-points of F, then

$$d(\overline{x},\overline{y}) = d(F(\overline{x}), F(\overline{y})) \leq C d(\overline{x},\overline{y}) \Rightarrow d(\overline{x},\overline{y}) = 0 \Leftrightarrow \overline{x} = \overline{y}$$

Task 1.4: Unique Zero with Banach-Fixed-Points

Define $\boldsymbol{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ by setting:

$$\boldsymbol{F}(x,y) = \begin{bmatrix} x + \frac{y}{2} \\ \frac{x}{2} + y + 1 \end{bmatrix}$$

Define the map $\boldsymbol{G}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{F}(\boldsymbol{x})$. Using the Banach-Fixed-Point theorem on $\boldsymbol{G}$, prove that $\boldsymbol{F}$ admits a unique zero, i.e. $\exists \overline{\boldsymbol{x}} \in \mathbb{R}^2 \colon \boldsymbol{F}(\overline{\boldsymbol{x}}) = \boldsymbol{0}$.

We begin by showing that $\boldsymbol{G}$ is Lipschitz-continuous with a Lipschitz-constant $C \in [0,1)$:

$$\begin{aligned}
&\|\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{G}(\boldsymbol{y})\| = \|\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{y} + \boldsymbol{F}(\boldsymbol{y})\| \\
&= \sqrt{\left(x_1 - x_1 - \frac{x_2}{2} - y_1 + y_1 + \frac{y_2}{2}\right)^2 + \left(x_2 - \frac{x_1}{2} - x_2 - 1 - y_2 + \frac{y_1}{2} + y_2 + 1\right)^2} \\
&= \sqrt{\left(\frac{y_2}{2} - \frac{x_2}{2}\right)^2 + \left(\frac{y_1}{2} - \frac{x_1}{2}\right)^2} = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\| = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|
\end{aligned}$$

Hence $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2 \colon \|\boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{G}(\boldsymbol{y})\| \leq \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|$, i.e. $\boldsymbol{G}$ is Lipschitz-continuous with Lipschitz-constant $C = \frac{1}{2}$. As $0 \leq C < 1$, G is a contraction on $\mathbb{R}^2$ under the euclidean norm, thus we can apply the Banach-Fixed-Point theorem, which states $\exists \overline{\boldsymbol{x}} \in \mathbb{R}^2 \colon \boldsymbol{G}(\overline{\boldsymbol{x}}) = \overline{\boldsymbol{x}}$. Hence:

$$\boldsymbol{G}(\overline{\boldsymbol{x}}) = \overline{\boldsymbol{x}} = \overline{\boldsymbol{x}} - \boldsymbol{F}(\overline{\boldsymbol{x}}) \Leftrightarrow \boldsymbol{0} = \boldsymbol{F}(\overline{\boldsymbol{x}})$$

Solving the equation yields the unique zero $\overline{\boldsymbol{x}}^T = \begin{bmatrix} \frac{4}{6} & -\frac{4}{3} \end{bmatrix}$.

Task 1.5: Local Invertibility

Define $\boldsymbol{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\boldsymbol{F}(x,y,z) = \begin{bmatrix} x + y + z \\ xy + yz + xz \\ xyz \end{bmatrix}$$

Determine all the points in $\mathbb{R}^3$ in which $\boldsymbol{F}$ is locally invertible.

Recall that $\boldsymbol{F}$ is locally invertible in $\boldsymbol{x}_0 \in \mathbb{R}^3$ if $\det\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}(\boldsymbol{x}_0)\right) \neq 0$, where $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}$ is the Jacobian of $\boldsymbol{F}$:

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ y+z & x+z & x+y \\ yz & xz & xy \end{bmatrix}$$

$$\det\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}\right) = (x+z)xy - xz(x+y) - (y+z)xy + yz(x+y) + (y+z)xz - yz(x+z)$$
$$= x^2y + xyz - (x^2z + xyz) - (xy^2 + xyz) + xyz + y^2z + xyz + xz^2 - xyz - yz^2$$
$$= x^2y - x^2z - xy^2 + y^2z + xz^2 - yz^2 = J(\boldsymbol{x})$$

One obvious non-invertible point is $\mathbf{0}$, as $J(\mathbf{0}) = 0$. Additionally, the lines $(\lambda, 0, 0)$, $(0, \xi, 0)$ and $(0, 0, \mu)$ provide subspaces, for $\lambda, \xi, \mu \in \mathbb{R}$, where $\boldsymbol{F}$ is not locally-invertible. Additionally we notice $J((\lambda, \lambda, \lambda)) = 0$, as the first and last row of $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}$ become linearly dependent.

Let $x = 0$ and $y, z \neq 0$, then $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}$ has the following form:

$$\begin{bmatrix} 1 & 1 & 1 \\ y+z & z & y \\ yz & 0 & 0 \end{bmatrix}$$

Setting $z = -y$ yields $J(\boldsymbol{x}) = -2y^3$, however for $z = y$ we get $J(\boldsymbol{x}) = 0$. Hence $\boldsymbol{F}$ is locally invertible on:

$$\mathbb{R}^3 \setminus (\operatorname{span}(\boldsymbol{e}_2, \boldsymbol{e}_3) \cup \operatorname{span}(\boldsymbol{e}_1, \boldsymbol{e}_3) \cup \operatorname{span}(\boldsymbol{e}_1, \boldsymbol{e}_2) \cup \operatorname{span}(\mathbf{1}_3))$$

where

$$\mathbf{1}_k = \sum_{i=1}^{k} \boldsymbol{e}_i \in \mathbb{R}^k$$