Recall on finding critical points. Lets $A \subseteq \mathbb{R}^n$ be open. Suppose $f\colon A \to \mathbb{R}$ is differentiable in $\boldsymbol{x}_0 \in A$. We say $\boldsymbol{x}_0$ is a critical point of f if $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$. Recall that a local minimum or maximum of f is always a critical point. Suppose $f \in \mathcal{C}^2$, then we define the Hessian-Matrix $\mathrm{H}f$ of f by:

$$\mathrm{H}f = \mathrm{J}\nabla f = \frac{\partial^2 f}{\partial x_i \partial x_j} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j=1,\dots,n}$$

Suppose $\boldsymbol{x}_0$ is a critical point of f, we categorize $\boldsymbol{x}_0$ by examining the definiteness of $\mathrm{H}f$:

- if $\mathrm{H}f(\boldsymbol{x}_0) > 0$ then $\boldsymbol{x}_0$ is a local minimum
- if $\mathrm{H}f(\boldsymbol{x}_0) < 0$ then $\boldsymbol{x}_0$ is a local maximum
- if $\mathrm{H}f(\boldsymbol{x}_0) \lessgtr 0$ then $\boldsymbol{x}_0$ is a saddle point

Task 1.1: Finding and classifying critical points

Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by setting

$$f(x, y) = \frac{xy}{1 + x^2 + y^2}$$

Find all critical points of f and classify them into local minima, maxima and saddle-points.

We first compute ∇f:

$$\nabla f = \begin{bmatrix} \frac{y(1+x^2+y^2)-2x^2y}{(1+x^2+y^2)^2} \\ \frac{x(1+x^2+y^2)-2xy^2}{(1+x^2+y^2)^2} \end{bmatrix} = \begin{bmatrix} \frac{y(1-x^2+y^2)}{(1+x^2+y^2)^2} \\ \frac{x(1-y^2+x^2)}{(1+x^2+y^2)^2} \end{bmatrix}$$

First we compute $\frac{\partial g^2}{\partial x}$ for $g(x, y) = 1 + x^2 + y^2$. Following the product rule yields:

$$\frac{\partial g^2}{\partial x} = 2\frac{\partial g}{\partial x} g = 4x(1 + x^2 + y^2)$$
$$\frac{\partial g^2}{\partial y} = 2\frac{\partial g}{\partial y} g = 4y(1 + x^2 + y^2)$$

Now we compute $\mathrm{H}f$ and introduce $h_1(x, y) = y(1 - x^2 + y^2)$ and $h_2(x, y) = h_1(y, x)$:

$$\frac{\partial \langle \boldsymbol{e}_1, \nabla f\rangle}{\partial x} = \frac{\partial}{\partial x}\frac{h_1}{g^2} = \frac{(\partial_x h_1)g^2 - h_1(\partial_x g^2)}{g^4}$$
$$\partial_x h_1(x, y) = -2xy \Rightarrow \frac{-2xyg^2 - 4xh_1 g}{g^4} = \frac{-2xg(yg + 2h_1)}{g^4} = \frac{-2x(yg + 2h_1)}{g^3}$$
$$\frac{\partial \langle \boldsymbol{e}_1, \nabla f\rangle}{\partial y} = \frac{\partial}{\partial y}\frac{h_1}{g^2} = \frac{(\partial_y h_1)g^2 - h_1(\partial_y g^2)}{g^4}$$
$$\partial_y h_1(x, y) = 1 - x^2 + 2y^2 \Rightarrow \frac{(1 - x^2 + 2y^2)g^2 - 4ygh_1}{g^4} = \frac{(1 - x^2 + 2y^2)g - 4yh_1}{g^3}$$
$$\frac{\partial \langle \boldsymbol{e}_1, \nabla f\rangle}{\partial x} = \frac{\partial}{\partial x}\frac{h_2}{g^2} = \frac{(\partial_x h_2)g^2 - h_2(\partial_x g^2)}{g^4}$$
$$\partial_x h_2(x, y) = 1 - y^2 + 2x^2 \Rightarrow \frac{(1 - y^2 + 2x^2)g^2 - 4xgh_2}{g^4} = \frac{(1 - y^2 + 2x^2)g - 4xh_2}{g^3}$$
$$\frac{\partial \langle \boldsymbol{e}_2, \nabla f\rangle}{\partial y} = \frac{\partial}{\partial y}\frac{h_2}{g^2} = \frac{(\partial_y h_2)g^2 - h_2(\partial_y g^2)}{g^4}$$
$$\partial_y h_2(x, y) = -2xy \Rightarrow \frac{-2xyg^2 - 4ygh_2}{g^4} = \frac{-2y(xg + 2h_2)}{g^3}$$

It follows that:

$$\mathrm{H}f = \begin{bmatrix} \frac{-2x(yg+2h_1)}{g^3} & \frac{(1-x^2+2y^2)g-4yh_1}{g^3} \\ \frac{(1-y^2+2x^2)g-4xh_2}{g^3} & \frac{-2y(xg+2h_2)}{g^3} \end{bmatrix}$$

Given $\mathrm{H}f$ and ∇f, we can find all critical points and classify them. One obvious solution where $\nabla f(\boldsymbol{x}) = \mathbf{0}$ is for $\boldsymbol{x} = \mathbf{0}$. However, since $\boldsymbol{x} \in \mathbb{R}^2$, we cannot find any other $\boldsymbol{x} \in \mathbb{R}^2$, such that $\nabla f(\boldsymbol{x}) = \mathbf{0}$. Next we compute $\mathrm{H}f(\mathbf{0})$:

$$\mathrm{H}f(\mathbf{0}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

If follows that $(\mathrm{H}f\mathbf{0})^* = \mathrm{H}f(\mathbf{0})$ thus we can just check it's eigenvalues for definiteness:

$$\det(\lambda \mathrm{I}_2 - \mathrm{H}f(\mathbf{0})) = \lambda^2 - 1 \stackrel{!}{=} 0 \Leftrightarrow \lambda = \pm 1$$

It follows that $\mathrm{H}f(\mathbf{0})$ is indefinite. Computing $\det \mathrm{H}f(\mathbf{0}) = -1$, we see that $\mathbf{0}$ is a saddle-point of f.

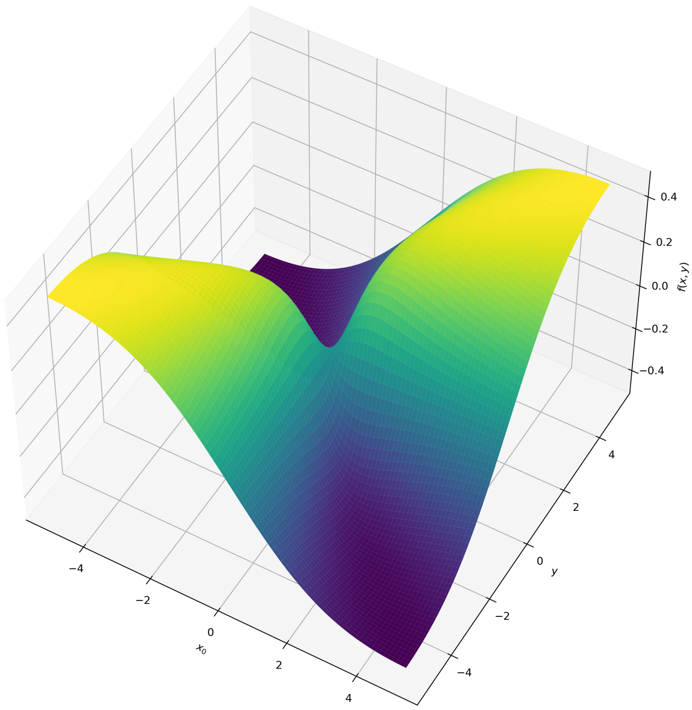

Figure 1: Surface-plot of f over $[-5, 5]^2$

Task 1.2: Finding and classifying critical points

Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by setting

$$f(x,y) = 2(x^4 + y^4 + 1) - (x+y)^2$$

Find all the critical points of f and classify them into local minima, maxima and saddle-points.

We first compute ∇f:

$$\frac{\partial f}{\partial x} = 8x^3 - 2x - 2y$$
$$\frac{\partial f}{\partial y} = 8y^3 - 2x - 2y$$
$$\mathrm{D}f = \begin{bmatrix} 8x^3 - 2x - 2y & 8y^3 - 2x - 2y \end{bmatrix}$$

We continue with $\mathrm{H}f$:

$$\frac{\partial \langle \boldsymbol{e}_1, \nabla f\rangle}{\partial x} = 24x^2 - 2 \qquad \frac{\partial \langle \boldsymbol{e}_1, \nabla f\rangle}{\partial y} = -2$$
$$\frac{\partial \langle \boldsymbol{e}_2, \nabla f\rangle}{\partial x} = -2 \qquad \frac{\partial \langle \boldsymbol{e}_2, \nabla f\rangle}{\partial y} = 24y^2 - 2$$

Thus

$$\mathrm{H}f = \begin{bmatrix} 24x^2 - 2 & -2 \\ -2 & 24y^2 - 2 \end{bmatrix} = (\mathrm{H}f)^T$$

Next we solve $\nabla f(\boldsymbol{x}) = \mathbf{0}$, resulting in the following system of equations.

$$8x^3 - 2x - 2y = 0$$
$$8y^3 - 2x - 2y = 0$$

Subtracting the two equations yields $x^3 = y^3$, thus since x^3 is injective, we get $x = y$, substituting in one equation yields

$$8x^3 - 4x = 0 \Leftrightarrow 4x(2x^2 - 1) = 0$$

This has one trivial solution $x = 0$ and the two remaining solutions $x = \pm\frac{\sqrt{2}}{2}$. Now we compute $\mathrm{H}f(x,x)$ and plug in our solutions:

$$\mathrm{H}f_1 = \mathrm{H}f\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \begin{bmatrix} \frac{48}{4} - 2 & -2 \\ -2 & \frac{48}{4} - 2 \end{bmatrix} = \begin{bmatrix} 20 & -2 \\ -2 & 20 \end{bmatrix}$$
$$\det(\lambda \mathrm{I}_2 - \mathrm{H}f_1) = (\lambda - 20)^2 - 4 = 0 \Leftrightarrow \lambda - 20 = \pm 2 \Leftrightarrow \lambda = 20 \pm 2$$

Thus $\mathrm{H}f_1 > 0$ and $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ is a local minimum.

$$\mathrm{H}f_2 = \mathrm{H}f\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \begin{bmatrix} 20 & -2 \\ -2 & 20 \end{bmatrix} = \mathrm{H}f_1$$

Thus $\mathrm{H}f_2$ is positive definite and $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ is a local minimum. For $\boldsymbol{x} = \mathbf{0}$, we compute $\det \mathrm{H}f(\mathbf{0}) = 0$, thus we cannot determine wether $\mathbf{0}$ is a saddle-point or not, by analyzing the hessian-matrix. Thus we limit f to one line $y = mx$ for $m \in \mathbb{R}$ and check the function of one variable

$$f_m(x) = f(x, mx) = 2(x^4 + m^4x^4 + 1) - (x(1+m))^2$$
$$= 2(x^4(1+m^4) + 1) - x^2(1+m)^2$$

Finding a local maximum of f_m can be done using the usual methods of real analysis in one dimension:

$$\frac{\mathrm{d}f_m}{\mathrm{d}x} = 8x^3(1+m^4) - 2x(1+m)^2 = 2x(x^2(1+m^4) - (1+m)^2)$$
$$= 2x(1+m^4)\left(x^2 - \frac{(1+m)^2}{(1+m^4)}\right)$$
$$\frac{\mathrm{d}^2 f_m}{\mathrm{d}x^2} = 24x^2(1+m^4) - 2(1+m)^2$$

We see that for $x = 0$ we get a local extremum, resulting in

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(0) = -2(1+m)^2$$

As $(1+m)^2 > 0 \forall m \in \mathbb{R}$, we get that $f''(0) < 0$ for all $m \in \mathbb{R}$, thus **0** is a local maximum.

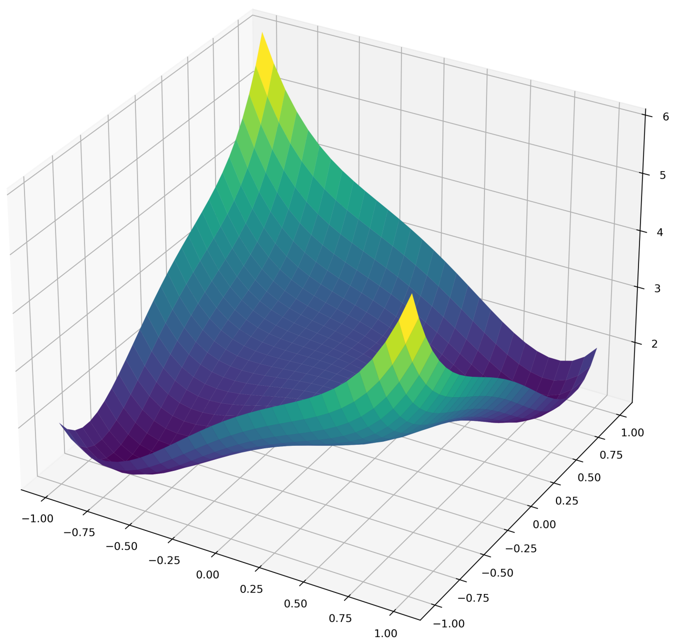

Figure 2: Plot of f over $[-1,1]^2$

Task 1.3: Global Critical Points on bound domain

Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by $f(x,y) = xy^2$ and consider the set

$$A = \{(x,y) \in \mathbb{R}^2 \colon y \geq 0 \wedge y \leq 1 + x \wedge y \leq 1 - x\}$$

Find the global maxima and minima of $f|_A$.

First we analyze A a bit by drawing it and determining the boundary ∂A and the open kernel $\mathring{A}$.

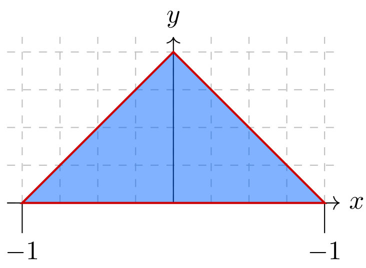

Figure 3: The set $A = B \cup I$

Where $B = \partial A$ and $I = \mathring{A}$. Computing ∇f yields:

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} y^2 \\ 2xy \end{bmatrix}$$

And $\mathrm{H}f$:

$$\mathrm{H}f = \begin{bmatrix} 0 & 2y \\ 2y & 2x \end{bmatrix}$$

We beginn by analyzing f on ∂A and set $y = x+1$ for $x \in [-1, 0]$, thus we get $f_1(x) = f(x, x+1) = x(x+1)^2 = x(x^2+2x+1) = x^3+2x^2+x$, yielding $f_1'(x) = 3x^2+4x+1$ and $f_1''(x) = 6x+2$. Finding the critical points of f_1 yields

$$f_1'(x) \overset{!}{=} 0 \Leftrightarrow x_{1,2} = \frac{-4 \pm \sqrt{16-12}}{6} = \frac{-4 \pm 2}{6}$$
$$\Rightarrow x_1 = -1 \qquad x_2 = -\frac{1}{3}$$

Thus $\boldsymbol{x}_1 = \begin{bmatrix} -1 & 0 \end{bmatrix}^T$ and $\boldsymbol{x}_2 = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \end{bmatrix}^T$ are critical points of f on ∂A. Next we set $f_2(x) = f(x, 1-x) = x(1-x)^2 = x(1-2x+x^2) = x^3-2x^2+x$, thus $f_2'(x) = 3x^2-4x+1$ and $f_2''(x) = 6x-4$. Solving $f_2'(x) = 0$ produces:

$$x_{1,2} = \frac{4 \pm \sqrt{16-12}}{6} = \frac{4 \pm 2}{6}$$
$$\Rightarrow x_1 = 1 \qquad x_2 = \frac{1}{3}$$

Thus $\boldsymbol{x}_3 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ and $\boldsymbol{x}_4 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}^T$ are critical points of f on ∂A. Since we are only interested in global extrema on A and $f(\boldsymbol{x}_2) < f(\boldsymbol{x}_1)$ and $f(\boldsymbol{x}_4) > f(\boldsymbol{x}_3)$, we see that although all points $(x, 0)$ are critical points of f, they do not form global extrema. Thus all global extrema of f on A are given by

$$\left\{ \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \right\}$$

Let $\boldsymbol{x} \in \mathbb{R}^2 \setminus \mathrm{span}(\boldsymbol{e}_1) = D$, then $f(\boldsymbol{x})$ has no critical points on D, therefore any open subset $C \subseteq D$ has no critical points of f, therefore $\mathring{A}$ includes no critical points f.

Task 1.4: Existence of critical points

a) Suppose $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\boldsymbol{x}_0 \in \mathring{A}$, where $\mathring{A}$ is the interior of A. Show that if $\boldsymbol{x}_0$ is a local minimum or maximum of f, then $\nabla f(\boldsymbol{x}_0) = \boldsymbol{0}$

b) (Rolle's Theorem in $\mathbb{R}^n$) Suppose that $A \subset \mathbb{R}^n$ is compact with $\mathring{A} \neq \emptyset$. Let $f \in \mathcal{C}(A)$, differentiable in $\mathring{A}$ and constant on ∂A. Prove there exists $\boldsymbol{x}_0 \in \mathring{A}$ such that $\nabla f(\boldsymbol{x}_0) = \boldsymbol{0}$.

We start with b). If f is constant on $\mathring{A}$, then $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathring{A}\colon f(\boldsymbol{x}) \geq f(\boldsymbol{y})$.

If f is not constant on $\mathring{A}$, then $\exists \boldsymbol{\xi} \in \mathring{A} \colon \forall \boldsymbol{x} \in \partial A \colon f(\boldsymbol{\xi}) \neq f(\boldsymbol{x})$. Since A is compact $\exists \boldsymbol{x} \in A$ such that $\forall \boldsymbol{w} \in A \colon f(\boldsymbol{x}) \geq f(\boldsymbol{w})$ following Weierstrass' theorem. Let, without loss of generality, $\boldsymbol{\xi} \in \mathring{A} \Rightarrow f(\boldsymbol{\xi}) > c$ where $f(\partial A) = \{c\}$. It follows

$$f(\boldsymbol{x}) \geq f(\boldsymbol{\xi}) > c$$

Hence $\boldsymbol{x} \notin \partial A \Leftrightarrow \boldsymbol{x} \in \mathring{A}$ as A is compact.

Now onto a). Let $\varepsilon > 0$ small enough that $\mathcal{B}_\varepsilon(\boldsymbol{x}_0) \subset A$, we define $l_k \colon \mathcal{B}_\varepsilon(0) \to \mathbb{R}^n$ with $l_k(t) = \boldsymbol{x}_0 + t\boldsymbol{e}_k$ As l_k is a linear-affine function on $\mathbb{R}^n$ it is differentiable for $\boldsymbol{x}_0 \in \mathbb{R}^n$, hence $f \circ l_k$ is differentiable in 0:

$$\mathrm{D}(f \circ l_k)(0) = \mathrm{D}f(l_k(0))\mathrm{D}l_k(0) = \mathrm{D}f(\boldsymbol{x}_0)\boldsymbol{e}_k = \frac{\partial f}{\partial x_k}(\boldsymbol{x}_0)$$

If $\forall k = 1, \ldots n \colon \frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) = 0$, then $\nabla f(\boldsymbol{x}_0)$.