NTNU
Kunnskap for en bedre verden

TTK4135 - Optimization and Control

Exercise #1

Author:
Sondre Pedersen

January 13, 2025

Problem 1: KKT example 1

Consider
$$\min x_1 + 2x_2 \qquad s.t. \qquad 2 - x_1^2 - x_2^2 \geq 0,\ x_2 \geq 0.$$

(a) Find the optimal point by inspecting the feasible region and the objective function.

(b) Check the KKT conditions at the optimal point.

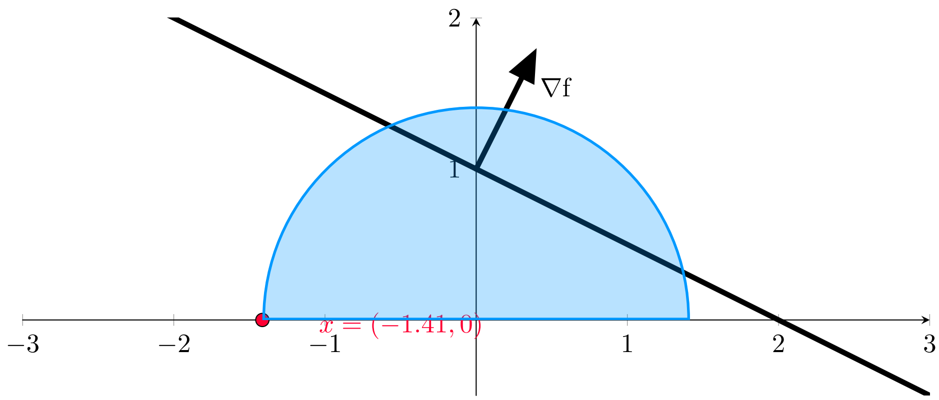
(c) Illustrate the gradients of the active constraints and the objective function at the solution.

(d) Explain why the Lagrange multipliers are positive.

(e) Is this problem a convex problem?

(a) Optimal point

Here is a figure showing the feasible region in blue, f and ∇f. From ∇f we can see that $x^* = \left[-\sqrt{2}, 0\right]^\top$ is the optimal point.



(b) KKT conditions

The KKT-conditions are in this case

$$\begin{aligned}
\nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= 0 &(1)\\
c_1(\mathbf{x}) &= 2 - x_1^2 - x_2^2 \geq 0 &(2)\\
c_2(\mathbf{x}) &= x_2 \geq 0 &(3)\\
\lambda &\geq 0 &(4)\\
\lambda_1 c_1(\mathbf{x}) &= 0 &(5)\\
\lambda_2 c_2(\mathbf{x}) &= 0 &(6)
\end{aligned}$$

.

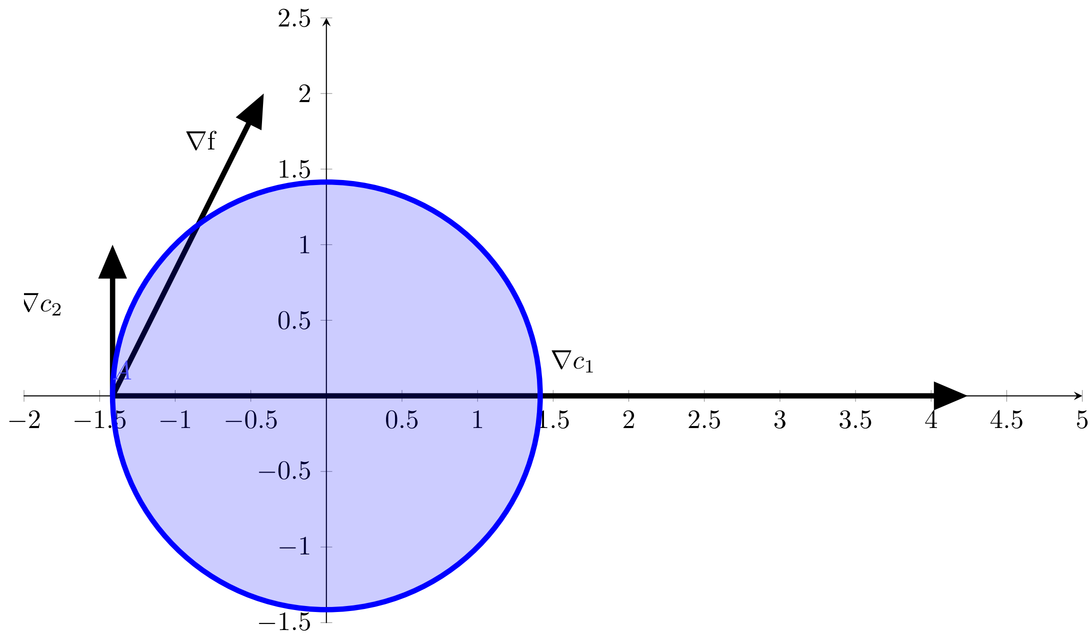
Use the Lagrangian to find λ:

$$\begin{aligned}
\mathcal{L}(\mathbf{x}, \lambda) &= x_1 + 2x_2 - \lambda_1(2 - x_1^2 - x_2^2) - \lambda_2(x_2)\\
\nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} 1 + 2x_1\lambda_1 \\ 1 + 2x_2\lambda_2 - \lambda_2 \end{bmatrix}\\
\nabla_x \mathcal{L}(\mathbf{x}^*, \lambda) &= \begin{bmatrix} 1 - 2\sqrt{2}\lambda_1 \\ 2 - \lambda_2 \end{bmatrix}\\
\implies \lambda^* &= \begin{bmatrix} \frac{1}{2\sqrt{2}} & 2 \end{bmatrix}^\top
\end{aligned}$$

This λ^* satisfies (1) and (4).
(2), (3), (5), and (6) are satisified since both conditions are binding ($c_1(\mathbf{x}^*) = c_2(\mathbf{x}^*) = 0$).

(c) Gradients

The gradients are $\nabla f(\mathbf{x}^*) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\nabla c_1(\mathbf{x}^*) = \begin{bmatrix} -2(-2\sqrt{2}) \\ -2(0) \end{bmatrix} = \begin{bmatrix} 4\sqrt{2} \\ 0 \end{bmatrix}$, and $\nabla c_2(\mathbf{x}^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Illustrated at $\mathbf{x}^* = \begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$:



(d) Lagrange multipliers

The lagrange multipliers are both positive, because the objective function would improve by should the constraints be less strict. The value of the multipliers indicate how much the objective function would improve should the restrictions be extended in the direction of their negative gradient.

(e) Convexity

This is a convex problem because (i) the objective function is linear (which means it is convex and concave), and (ii) because the feasible region is convex. The convexity of the feasible region is easy to see, and can also be show by finding the hessian for c_1:

$$\nabla^2_{xx}\mathcal{L}(\mathbf{x}, \lambda^*) = \begin{bmatrix} 2\lambda_1 & 0 \\ 0 & 2\lambda_2 \end{bmatrix}.$$

Since $\lambda \geq 0$, the Hessian is positive semi-definite, which means the region is convex. $c_1(\mathbf{x})$ is also convex because it is linear.

Problem 2: KKT example 2

Consider

$$\min 2x_1 + x_2 \qquad s.t. \qquad x_1^2 + x_2^2 - 2 = 0.$$

(a) Find all points which satisfy the KKT conditions

(b) Illustrate the gradients of the active constraint and the objective function at the KKT points.

(c) What is the value of the Lagrange multiplier? Is this consistent with the KKT conditions?

(d) Explain why each KKT point is (or is not) a local solution to the optimization problem.

(e) Check the 2nd order conditions for the KKT points.

(f) Is this problem a convex problem?

(a) KKT points

The KKT points must satisfy $\nabla_x \mathcal{L}(\mathbf{x}, \lambda) = 0$.

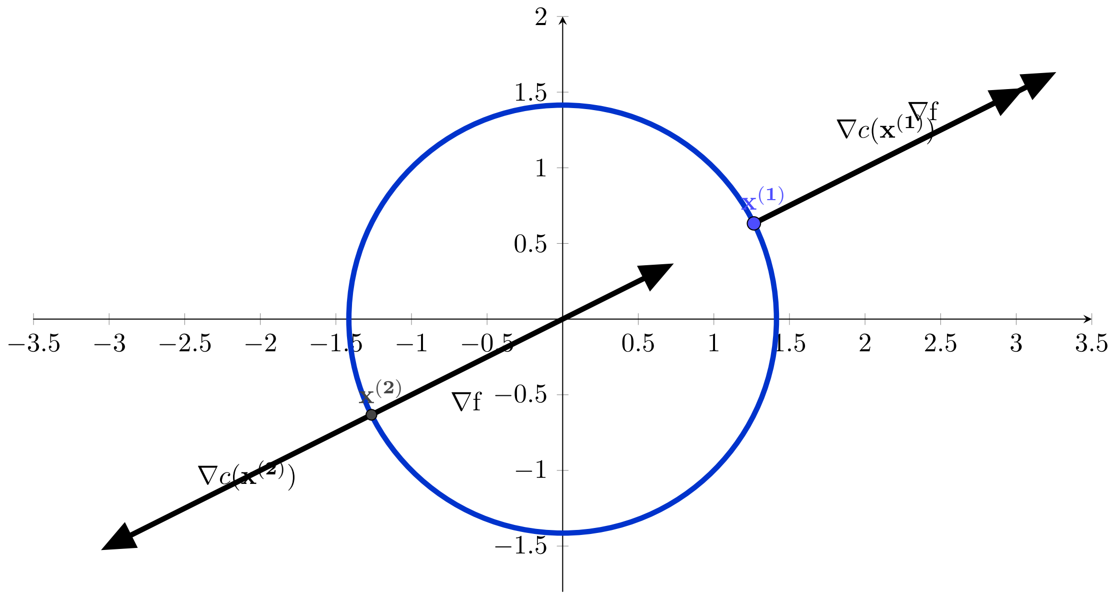
$$
\begin{aligned}
\mathcal{L}(\mathbf{x},\lambda) &= 2x_1 + x_2 - \lambda(x_1^2 + x_2^2 - 2)\\
\nabla_x \mathcal{L}(\mathbf{x},\lambda) &= \begin{bmatrix} 2-2\lambda x_1 \\ 1-2\lambda x_2 \end{bmatrix} = 0\\
\implies \lambda &= \frac{1}{x_1} = \frac{1}{2x_2}\\
\implies x_2 &= \pm\sqrt{\frac{2}{5}},\ x_1 = 2x_2 = \pm\sqrt{\frac{8}{5}}
\end{aligned}
$$

The KKT points are $\mathbf{x^{(1)}} = \begin{bmatrix}\sqrt{\frac{8}{5}} & \sqrt{\frac{2}{5}}\end{bmatrix}^\top$ and $\mathbf{x^{(2)}} = \begin{bmatrix}-\sqrt{\frac{8}{5}} & -\sqrt{\frac{2}{5}}\end{bmatrix}^\top$

(b) Gradients

The gradiants are

$$
\begin{aligned}
\nabla_x f(\mathbf{x}) = \begin{bmatrix} 2 \\ 1 \end{bmatrix},\ \nabla c_1(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}\\
\nabla c_1(\mathbf{x^{(1)}}) = \begin{bmatrix} \sqrt{\frac{16}{5}} \\ \sqrt{\frac{4}{5}} \end{bmatrix},\ \nabla c_1(\mathbf{x^{(2)}}) = \begin{bmatrix} -\sqrt{\frac{16}{5}} \\ -\sqrt{\frac{4}{5}} \end{bmatrix}
\end{aligned}
$$



(c) Lagrange multipliers

We see that $\lambda^{(1)} = \frac{1}{x_1} = \frac{1}{\sqrt{\frac{8}{5}}} = 0.79$ and $\lambda^{(2)} = -0.79$. The sign of the lagrange multiplier does however not matter as much in this problem because of the equality constraint. If we solved the equivalent problem with $c_1^{'}(\mathbf{x}) = -c_1(\mathbf{x})$, the lagrange multiplier would have changed sign for the two points.

(d) The solution

$\mathbf{x^{(1)}}$ maximises f. We can see this from the illustration, where the constraint gradient points in the same direction as the objective function gradient. $\mathbf{x^{(2)}}$ is a local solution to the problem, since the constraint gradient is opposite to the objective gradient.

(e) 2nd order conditions

For this we need the Hessian. $\nabla^2_{xx}\mathcal{L}(\mathbf{x},\lambda) = \begin{bmatrix}-2\lambda & 0\\ 0 & 2\lambda\end{bmatrix}$. We see that $\nabla^2_{xx}\mathcal{L}(\mathbf{x^{(1)}},\lambda^{(1)}) = \begin{bmatrix}-\sqrt{\frac{16}{5}} & 0\\ 0 & -\sqrt{\frac{4}{5}}\end{bmatrix}$ is negative semi-definite. This means that $\mathbf{x^{(1)}}$ is a local maximum. $\nabla^2_{xx}\mathcal{L}(\mathbf{x^{(2)}},\lambda^{(2)}) = \begin{bmatrix}\sqrt{\frac{16}{5}} & 0\\ 0 & \sqrt{\frac{4}{5}}\end{bmatrix}$ is positive semi-definite, a local minimum.

Here, $d^\top \nabla^2_{xx}\mathcal{L}(\mathbf{x^{(2)}},\lambda^{(2)})d > 0,\ d > 0$

(f) Convexity

This is not a convex problem, because the feasible region is not convex. A line between any distinct points would leave the feasible region.

Problem 3: Second-order conditions

Consider

$$\min_{x\in\mathbb{R}^2} f(x) = -2x_1 + x_2 \qquad s.t. \qquad \begin{cases} c_1(x) = (1-x_1)^3 - x_2 \geq 0 \\ c_2(x) = x_2 + 0.25x_1^2 - 1 \geq 0 \end{cases}$$

The optimal solution is $x^* = (0,1)^\top$, where both constraints are active.

(a) Do the LICQ hold at this point?

(b) Are the KKT conditions satisfied?

(c) Write down the sets $\mathcal{F}(x^*)$ and $\mathcal{C}(x^*,\lambda^*)$.

(d) Are the second-order necessary conditions satisifed?

(a) LICQ conditions

To check if the LICQ conditions hold, we see if the constraint gradients are linearly independent.

$$\nabla c_1(\mathbf{x}) = \begin{bmatrix}3(1-x_1)^2(-1)\\ -1\end{bmatrix}$$
$$\nabla c_2(\mathbf{x}) = \begin{bmatrix}0.5x_1\\ 1\end{bmatrix}$$
$$rank\left(\begin{bmatrix}\nabla c_1(\mathbf{x^*}^\top) & \nabla c_2(\mathbf{x}^*)^\top\end{bmatrix}\right) = rank\left(\begin{bmatrix}-3 & -1\\ 0 & 1\end{bmatrix}\right) = 2$$

which means that LICQ conditions are satisifed.

(b) KKT

Yes, they are satisfied with $\lambda^* = \begin{bmatrix}\frac{2}{3} & \frac{5}{3}\end{bmatrix}^\top$:

$$\mathcal{L}(\mathbf{x},\lambda) = -2x_1 + x_2 - \lambda_1((1-x_2)^3 - x_2) - \lambda_2(x_2 + 0.25x_1^2 - 1)$$
$$\nabla_x\mathcal{L}(\mathbf{x}^*,\lambda^*) = \begin{bmatrix}-2 + \lambda_1^* - 3(1-x_1^*)^2 - 0.5\lambda_2^* x_1^*\\ 1 + \lambda_1^* - \lambda_2^*\end{bmatrix}$$
$$= \begin{bmatrix}-2 + \frac{2}{3}\times 3(1-0)^3\\ 1 + \frac{2}{3} - \frac{5}{3}\end{bmatrix} = \underline{0}$$
$$c_1(\mathbf{x}^*) = c_2(\mathbf{x}^*) = 0 \implies \begin{cases} c_1(\mathbf{x}^*) \geq 0\\ c_2(\mathbf{x}^*) \geq 0\\ \lambda_1^* c_1(\mathbf{x}^*) = 0\\ \lambda_2^* c_2(\mathbf{x}^*) = 0\end{cases}$$

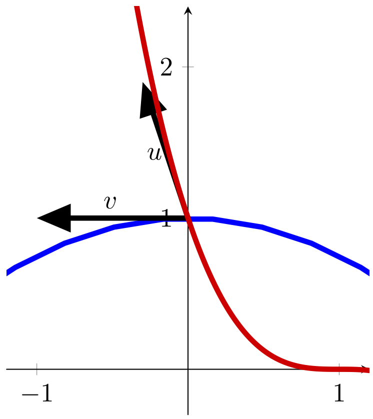
And of course, $\lambda \geq 0$, satisfying all the KKT conditions.

(c) Directions

Not really sure how to do this in general, but finding tangents to the constraint gradients was my strategy here. $\nabla c_1(\mathbf{x}^*) = \begin{bmatrix}-1 & -1\end{bmatrix}^\top$ and $\nabla c_2(x^*) = \begin{bmatrix}0 & 1\end{bmatrix}^\top$. The tangents are $\begin{bmatrix}-1 & 0\end{bmatrix}^\top$ and $\begin{bmatrix}-1 & 3\end{bmatrix}^\top$. Looking at a drawing was the only reasonable way to find them. Anyway here are all the directions between these constraint tangents:

$$\mathcal{F}(x^*) = \left\{ t\alpha \begin{bmatrix}-1\\3\end{bmatrix} + t(1-\alpha)\begin{bmatrix}-1\\0\end{bmatrix} \right\},\ \alpha \in [0,1],\ t \geq 0.$$

$\mathcal{F}$ creates a convex combination of the vectors u and v in the figure below.



To find $\mathcal{C}(\mathbf{x}^*)$, we need to find $\mathbf{d}$ such that $\nabla c_1(\mathbf{x}^*)^\top \mathbf{d} = 0$ and $\nabla c_2(\mathbf{x}^*)^\top \mathbf{d} = 0$.
We have the gradients from earlier, so

$$\begin{aligned}
\begin{bmatrix}-3 & -1\end{bmatrix}\begin{bmatrix}d_1\\d_2\end{bmatrix} &= 0\\
\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}d_1\\d_2\end{bmatrix} &= 0\\
\implies \mathbf{d} &= \mathbf{0}\\
\mathcal{C}(\mathbf{x}^*) &= \mathbf{0}
\end{aligned}$$

(d) Second order conditions

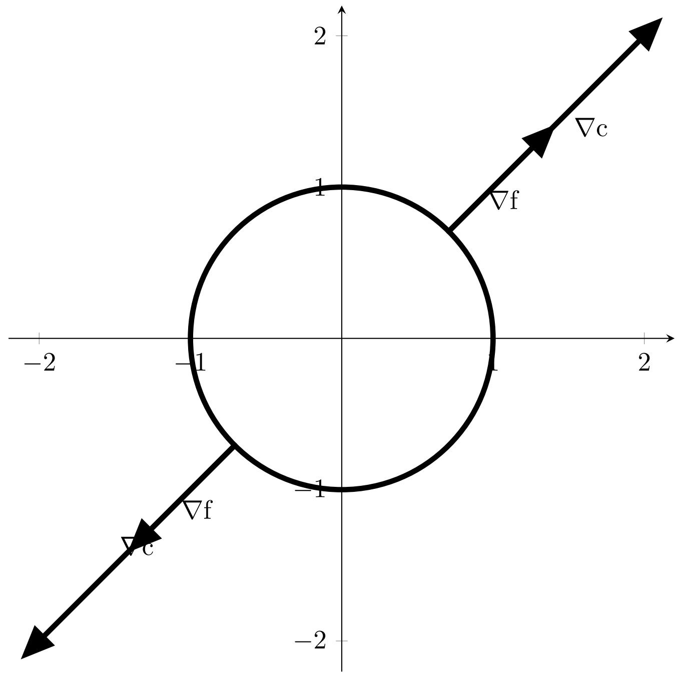
Yes, the necessary condition is satisifed, $\mathbf{d}^\top \nabla^2_{xx}\mathcal{L}(\mathbf{x}^*, \lambda^*)\mathbf{d} = \begin{bmatrix}0 & 0\end{bmatrix} \nabla^2_{xx}\mathcal{L}(\mathbf{x}^*, \lambda^*)\begin{bmatrix}0\\0\end{bmatrix} = 0$.

The sufficient condition is also satisfied, because there are no $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*)$ where $\mathbf{d} > \mathbf{0}$

Problem 4: Find optimizers

Find the *maxima* of the function $f(x) = x_1x_2$ over the unit disc defined by the inequality constraint $x_1^2 + x_2^2 = 1$. Illustrate the gradients of the active constraint and the objective function at the optimal points.

$$\begin{aligned}
\mathcal{L}(\mathbf{x}, \lambda) &= x_1x_2 - \lambda_1(x_1^2 + x_2^2 - 1)\\
\nabla\mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} x_2 - 2\lambda x_1\\ x_1 - 2\lambda x_2\\ -x_1^2 - x_2^2 + 1\end{bmatrix} = 0\\
\implies x_1 = x_2 &= \pm\frac{1}{\sqrt{2}},\ \lambda = \frac{1}{2}
\end{aligned}$$


2
1
−1
−2
−2
−1
1
2
∇c
∇f
∇f
∇c