1. Quadratic inequalities

Consider the set $S$ defined by the following inequalities:

$$(x_1 \geq -x_2 + 1 \text{ and } x_1 \leq 0) \text{ or } (x_1 \leq -x_2 + 1 \text{ and } x_1 \geq 0).$$

To be more precise,

$$S_1 = \{ \vec{x} \in \mathbb{R}^2 \mid x_1 \geq -x_2 + 1, x_1 \leq 0 \}$$
$$S_2 = \{ \vec{x} \in \mathbb{R}^2 \mid x_1 \leq -x_2 + 1, x_1 \geq 0 \}$$

$$S = S_1 \cup S_2.$$

(a) Draw the set $S$. Is it convex?

Solution:

![Figure 1: Set S](image)

The set $S$ as shown in Fig. 1 is not convex. We can prove this by counterexample. $(0, 2)$ and $(1, 0)$ both belong to the set, but the midpoint $(1/2, 1)$ does not.

(b) Show that the set $S$, can be described as a single quadratic inequality of the form

$$q(\vec{x}) = \vec{x}^\top A \vec{x} + 2\vec{b}^\top \vec{x} + c \leq 0,$$

for matrix $A = A^\top \in \mathbb{R}^{2 \times 2}$, $\vec{b} \in \mathbb{R}^2$ and $c \in \mathbb{R}$ i.e $S$ can be written as $S = \{ \vec{x} \in \mathbb{R}^2 \mid q(\vec{x}) \leq 0 \}$. Find $A, \vec{b}, c$.

Hint: Can you combine the constraints to make one quadratic constraint?
Solution: Within set $S$, $x_1 + x_2 - 1 \geq 0$ when $x_1 \leq 0$ and $x_1 + x_2 - 1 \leq 0$ when $x_1 \geq 0$. It follows that $q(x) = x_1(x_1 + x_2 - 1) \leq 0$ if and only if it is in the set. Expressing $q(x)$ in the desired form:

$$q(x) = x_1^2 + x_1x_2 - x_1 = x^T A x + 2b^T x + c$$

where

$$A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} -1/2 \\ 0 \end{bmatrix}, \quad c = 0.$$

(c) Give the definition of the convex hull of a set. What is the convex hull of this set, i.e., $S$?

Solution: The convex hull of the set is the whole space, $\mathbb{R}^2$. To see this note than any point $z = (z_1, z_2) \in \mathbb{R}^2$ can be written as $z = \frac{x+y}{2}$ with $x, y \in S$ as follows:

$$x = (2z_1, 1 - 2z_1), y = (0, 2(z_1 + z_2) - 1).$$

(d) We will now consider some convex optimization problems over $S_1$ that illustrate the role of the constraints in the optimization problem. For each of the following optimization problems find the optimal point, $\mathbf{x}^*$. Describe the constraints that are active in attaining the optimal value. Hint: Suppose that there exists a point $\mathbf{x}$ such that $\nabla f(\mathbf{x}) = 0$. From the first order characterization of a convex function $\mathbf{x}$ would be an optimum value for $f$ subject to no constraints. If $\mathbf{x}$ is not in the constraint set $S_1$, then the optimum point must be on the boundary of the set, i.e., it satisfies at least one of the constraints defining $S_1$ with equality.

i. Minimize $f(\mathbf{x}) = (x_1 + 1)^2 + (x_2 - 3)^2$ subject to $\mathbf{x} \in S_1$.

Solution: We first compute the unconstrained optimal value of $f$. Notice that $f$ is a convex function. Therefore, we can compute its optimal value by computing its gradient and setting it to 0. Doing so, we obtain the optimal value of $f$ to be 0 attained at the point $\mathbf{x}^* = (-1, 3)$. Now, since $\mathbf{x}^* \in S_1$, $\mathbf{x}^*$ is the solution to the constrained optimization problem as well.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{This figure illustrates the position of the optimum, $x^* = (-1, 3)$, and the level sets of the objective function, $f$, which are concentric circles around $x^*$.}
\end{figure}
ii. Minimize $f(\vec{x}) = (x_1 + 2)^2 + (x_2 - 2)^2$ subject to $\vec{x} \in S_1$.

**Solution:**

Proceeding as in the proof for the previous problem, we first find the solution to the unconstrained optimization problem. We get that the unconstrained problem is minimized at the point $\vec{x}_u^* = (-2, 2)$. However, this point is not in the feasible set, $S_1$. Therefore, the true optimum, $\vec{x}^*$, has one or more constraints active. Now, we will attempt to solve the problem with one active constraint. Suppose the one active constraint is $x_1 \geq -x_2 + 1$. Since this constraint is active, we must try and minimize $f(\vec{x})$ subject to $\vec{x}$ satisfying $x_1 = -x_2 + 1$.

Note that any point on this line can be written in the form $(0, 1) + \alpha(-1, 1)$. Now consider the function, $g(\alpha)$:

$$g(\alpha) = f((0, 1) + \alpha(-1, 1)) = (\alpha - 2)^2 + (\alpha - 1)^2.$$ 

Note that the function, $f(\alpha)$, is convex in $\alpha$. Therefore, we can minimize $g(\alpha)$ by taking its derivative and setting it to 0. By doing this, we get that $\alpha = 3/2$ is the unique minimizer of $g(\alpha)$. Therefore, the minimizer of $f$ subject to $x_1 = -x_2 + 1$ is the point $(-3/2, 5/2)$, and the function value is 0.5. Similarly, the minimizer of $f$ assuming the second constraint, $x_1 \leq 0$, is active is obtained at the point $(0, 2)$, and the function value at this point is 4, which is higher than the value at $(-3/2, 5/2)$. The final possibility is that both constraints are active. However, the optimal value of $f$ subject to both constraints being active will be greater than the value of $f$ obtained at $(-3/2, 5/2)$ which is in $S_1$. Therefore, we get that $f(\vec{x})$ is minimized at the point $\vec{x}^* = (-3/2, 5/2)$ subject to $\vec{x} \in S_1$. There is one active constraint at $\vec{x}^*$.

![Figure 3](image-url)  

*Figure 3:* This figure illustrates the position of the optimum, $x^* = (-3/2, 5/2)$, and the level sets of the objective function, $f$, which are concentric circles around $(-2, 2)$. Note that in this case, the unconstrained optimum does not lie in the set, $S_1$ and the optimal point lies on the boundary of one of the constraints.

iii. Minimize $f(\vec{x}) = x_1^2 + x_2^2$ subject to $\vec{x} \in S_1$.

**Solution:** Proceeding as before, we first check the case where 0 constraints are active. However, the unconstrained minimizer of $f$ is $(0, 0)$ which is not in $S_1$. Now, we check the cases
where one of the constraints is active. Assume that the constraint $x_1 \leq 0$ is active. In this case the optimizer is again obtained at the point $(0, 0)$ which is not in $S_1$. We then consider the case where the constraint $x_1 \geq -x_2 + 1$ is active. As before, we define the function $g(\alpha)$ as:

$$g(\alpha) = f((0, 1) + \alpha(-1, 1)) = \alpha^2 + (\alpha + 1)^2.$$ 

By optimizing over $\alpha$ by setting its gradient with respect to $\alpha$ and setting it to 0, we get the optimal setting of $\alpha$ is $-1/2$. However, note that the point $(1/2, 1/2)$ does not belong to $S_1$ either. Therefore, the only remaining possibility is the possibility that both constraints are active. This can happen solely at the point $(0, 1)$. At this point, the value of the function $f$ is 1, the optimizer $\hat{x}^* = (0, 1)$ and both constraints are active at $\hat{x}^*$.

Figure 4: This figure illustrates the position of the optimum, $x^* = (0, 1)$, and the level sets of the objective function, $f$, which are concentric circles around $(0, 0)$. Note that in this case, the unconstrained optimum does not lie in the set, $S_1$ and the optimal point lies on the boundary of both of the constraints.
2. About general optimization

In this exercise, we test your understanding of the general framework of optimization and its language.

We consider an optimization problem in standard form:

\[ p^* = \min_{x \in \mathbb{R}^n} f_0(x) : f_i(x) \leq 0, \ i = 1, \ldots, m. \]

In the following we denote by \( \mathcal{X} \) the feasible set. Note that the feasible set is a subset of \( \mathbb{R}^n \) that satisfies the inequalities \( f_i(x) \leq 0, \ i = 1, \ldots, m \). We make no assumption about the convexity of \( f_0(x) \) and \( f_i(x), \ i = 1, \ldots, m \). For the following statements, provide a proof or counter-example.

(a) A general optimization problem can be expressed as one with a linear objective.

**Solution:** The statement is true:

\[ p^* = \min_{x \in \mathcal{X}, t} t : t \geq f_0(x). \]

(b) A general optimization problem can be expressed as one without any constraints.

**Solution:** Again the statement is true: let us define

\[ g(x) := \begin{cases} f_0(x) & \text{if } x \in \mathcal{X}, \\ +\infty & \text{otherwise.} \end{cases} \]

Then

\[ p^* = \min_x g(x). \]

(c) A general optimization problem can be recast as a linear program (minimizing a linear objective subject to linear constraints), provided one allows for an infinitely many constraints.

**Solution:** This is true again; we have

\[ p^* = \max_t t : t \leq f_0(x) \text{ for every } x \in \mathcal{X}. \]

Alternatively,

\[ p^* = -\min_t -t : t \leq f_0(x) \text{ for every } x \in \mathcal{X}. \]

(d) If any of the constraint inequalities is strict (and therefore not active) at the optimum point, then we can remove the constraint from the original problem and obtain the same optimum value.

**Note:** Review the definition of active constraints from the textbooks: Boyd Section 4.1.1 and El Ghaoui Section 8.3.

**Hint:** Consider the problem

\[ \min_x f(x) = \begin{cases} x^2 & \text{if } |x| \leq 1, \\ -1 & \text{otherwise.} \end{cases} \]

such that \( |x| \leq 1 \)

**Solution:** This is not true in general. Consider the problem

\[ p^* := \min_x f_0(x) : |x| \leq 1, \]
where

\[ f_0(x) = \begin{cases} 
    x^2 & \text{if } |x| \leq 1, \\
    -1 & \text{otherwise.}
\end{cases} \]

The constraint \(|x| \leq 1\) is not active at the optimum \(x^* = 0\), and \(p^* = 0\). However, if we remove it, the new optimal value becomes \(-1\).
3. Trust region

In optimization a trust region refers to the region where a certain model (usually quadratic) can be used to approximate the original objective function.

Consider the problem

\[ p^* = \min_x x^\top Qx + 2c^\top x : \|x\|_2 = 1. \]

where \( Q \in S^n \) is symmetric (not necessarily positive semi-definite) and \( c \in \mathbb{R}^n \).

(a) Is the problem, as stated convex? What if \( Q \) is positive semi-definite?

**Solution:** The problem, as stated, is not convex, due to the non-convex constraint. This is true even when \( Q \) is positive semi-definite; in that case only the objective is convex.

(b) Show that the problem can be reduced to

\[ p^* = \min_y \sum_{i=1}^n (\lambda_i y_i^2 + 2d_i y_i) : \sum_{i=1}^n y_i^2 = 1, \]

for appropriate vectors \( \lambda, d \in \mathbb{R}^n \), which you will determine as functions of the problem data.

**Solution:** Let \( Q = U \Lambda U^\top \) be the eigenvalue decomposition of \( Q \), with \( U \) orthonormal and \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) the diagonal matrix containing the eigenvalues in decreasing order. With the change of variable \( y = U^\top x \), and with \( d = U^\top c \), we have \( c^\top x = d^\top y \), and \( x^\top Qx = y^\top \Lambda y \). Since \( U \) is orthonormal the constraint on \( x \) becomes \( \|y\|_2 = 1 \). This proves the result.

(c) Show that the problem can be further reduced to the convex problem

\[ p^* = \min_z \sum_{i=1}^n (\lambda_i z_i - 2|d_i|\sqrt{z_i}) : \sum_{i=1}^n z_i = 1, \quad z \geq 0. \]

**Solution:** Note that given a solution \( y^* \), if \( y_i \) is the same sign as \( d_i \), you can decrease the objective value by changing the sign of \( y_i \) while still remaining feasible. Therefore, at optimum, \( y_i \) must have the opposite sign of \( d_i \). Then the problem can be written as

\[ \min_\xi \sum_{i=1}^n (\lambda_i \xi_i^2 - 2|d_i|\xi_i) : \sum_{i=1}^n \xi_i^2 = 1, \quad \xi \geq 0, \]

with \( \xi_i = -\text{sign}(d_i)y_i, \quad i = 1, \ldots, n. \)

The new formulation results from the change of variable \( z_i = \xi_i^2, \quad i = 1, \ldots, n. \) The optimal \( y \) is then obtained as \( y_i = -\text{sign}(d_i)\sqrt{z_i}, \quad i = 1, \ldots, n. \)
4. (OPTIONAL) Strictly Diagonally Dominant Matrices

(a) Given a symmetric matrix \( A \in \mathbb{R}^{n \times n} \) we say that \( A \) is diagonally dominant if

\[
\forall i \in \{1, \ldots, n\} \text{ we have } A_{i,i} \geq \sum_{j \neq i} |A_{i,j}|
\]

That is, the diagonal entry is greater than the sum of absolute values of off diagonal entries of that row (equivalently column, as it’s symmetric). Prove that \( A \) is positive semi-definite.

**Solution:** To prove that \( A \) is PSD, we have to show that for any \( \vec{x} \in \mathbb{R}^n \), \( \vec{x}^\top Ax \geq 0 \).

\[
x^\top Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} x_i x_j
= \sum_{i=1}^{n} x_i^2 A_{i,i} + \sum_{j \neq i} x_i x_j A_{i,j}
\geq \sum_{i=1}^{n} x_i^2 A_{i,i} - \sum_{j \neq i} |x_i| \cdot |x_j| \cdot |A_{i,j}|
\geq \sum_{i=1}^{n} x_i^2 \sum_{j \neq i} |A_{i,j}| - \sum_{j \neq i} |x_i| \cdot |x_j| \cdot |A_{i,j}|
= \sum_{i=1}^{n} \sum_{j > i} [A_{i,j} (x_i^2 + x_j^2 - 2|x_i| |x_j|)]
= \sum_{i=1}^{n} \sum_{j > i} |A_{i,j}||x_i| |x_j|^2
\geq 0
\]

(b) Show that \( f(x) = \log \left( \sum_{i=1}^{n} e^{x_i} \right) \) is a convex function. You might find the previous part helpful.

**Solution:** Computing the Hessian of \( f \),

\[
\frac{\partial f(x)}{\partial x_i} = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}
\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{e^{x_i} \sum_{j=1}^{n} e^{x_j} - e^{2x_i}}{\left( \sum_{j=1}^{n} e^{x_j} \right)^2}
\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = -\frac{e^{x_i + x_j}}{\left( \sum_{j=1}^{n} e^{x_j} \right)^2}
\]

This is clearly symmetric and diagonally dominant and hence is PSD which implies \( f \) is convex.
5. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

*NOTE:* If you didn’t work with anyone, you can put “none” as your answer.