Self grades are due at 11 PM on November 3, 2023.

1. Simple Constrained Optimization Problem

Consider the optimization problem

$$\min_{x_1,x_2\in\mathbb{R}} \quad f(x_1,x_2) \tag{1}$$
$$\text{s.t.} \quad 2x_1 + x_2 \geq 1, \tag{2}$$
$$x_1 + 3x_2 \geq 1, \tag{3}$$
$$x_1 \geq 0,\ x_2 \geq 0. \tag{4}$$

(a) Make a sketch of the feasible set.

Solution: See Figure 1.

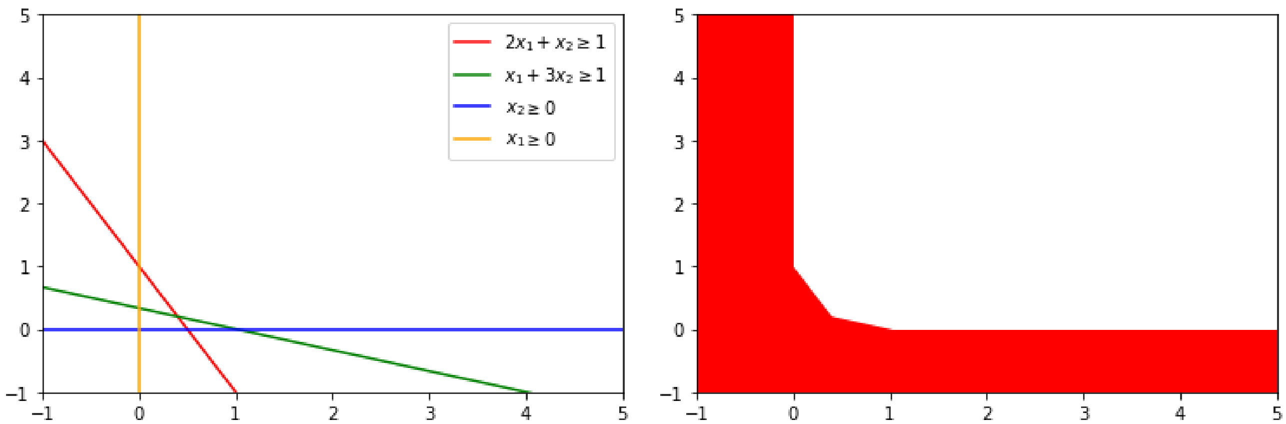


Figure 1: The feasible set is in white on the right figure.

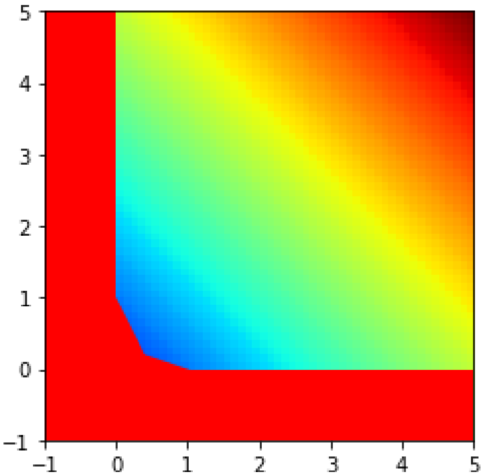
For each of the following objective functions, give the optimizer or set of optimizers, as well as the corresponding optimal value.

(b) $f(x_1,x_2) = x_1 + x_2$.

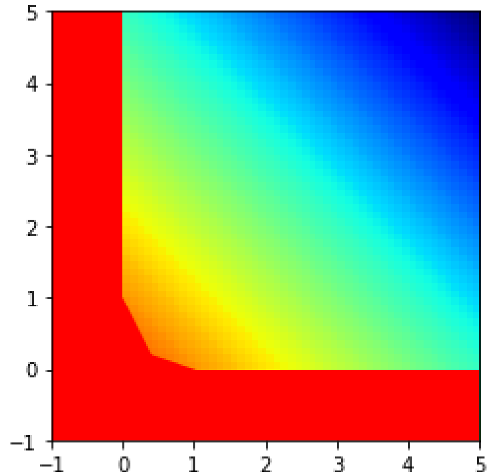
Solution: Using the drawing (Figure 2) it seems that the solution is such that $x_1^\star = \frac{2}{5}$ and $x_2^\star = \frac{1}{5}$. This can be seen by observing that $x_1 + x_2 \geq \frac{3}{5}$ for all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in the feasible set, with equality at $\begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \end{bmatrix}$. One can also verify the optimality of such point using the first order convexity condition:

$$\nabla f\left(\frac{2}{5},\frac{1}{5}\right)^\top \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \end{bmatrix}\right) \geq 0, \quad \forall (x_1,x_2) \in \mathcal{X}, \tag{5}$$

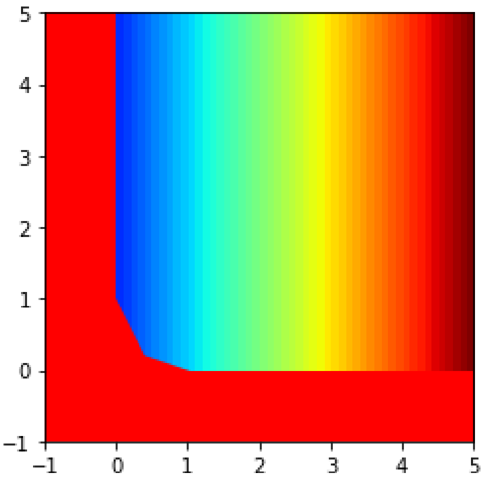
where $\mathcal{X}$ is the feasible set. The corresponding optimal value is thus $\frac{3}{5}$.



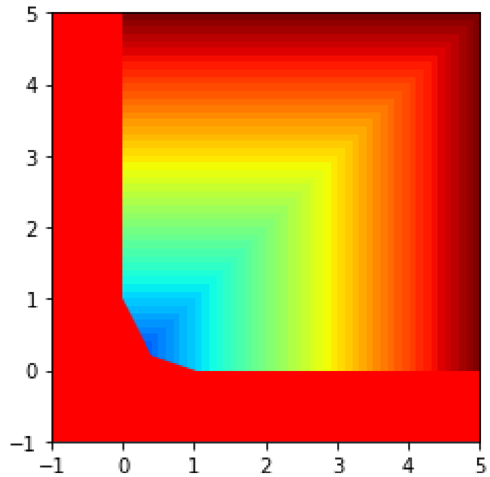
(b)



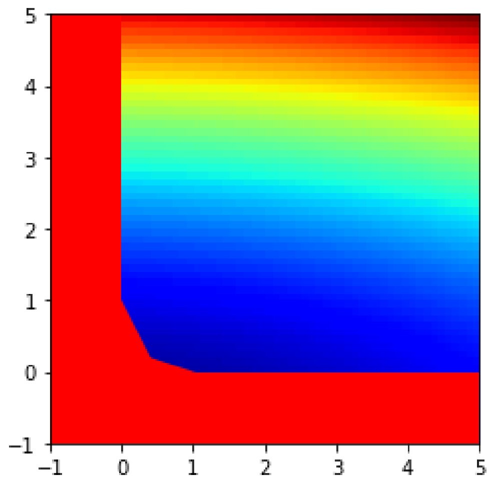
(c)



(d)



(e)



(f)

Figure 2: Solution of (b): $\vec{x}^\star = (\frac{2}{5}, \frac{1}{5})$, (c) is unbounded below, solutions of (d): $\vec{x}^\star = \{(0, x_2) \mid x_2 \geq 1\}$, solution of (e): $\vec{x}^\star = (\frac{1}{4}, \frac{1}{4})$, solution of (f): $\vec{x}^\star = (\frac{1}{2}, \frac{1}{6})$. In red is the infeasible points, then the level sets are shown with colors; blue points are points (x_1, x_2) with the lowest value $f(x_1, x_2)$, red points are the ones with highest value.

(c) $f(x_1, x_2) = -x_1 - x_2$.

Solution: Here (Figure 2) the problem is unbounded below as if $(x_1, x_2) = t(1, 1)$ with $t \geq 0$ then (x_1, x_2) is always feasible and $-2t \to -\infty$ when $t \to \infty$. The optimal value is thus $-\infty$.

(d) $f(x_1, x_2) = x_1$.

Solution: The optimal set is $S = \{\vec{x} | x_1 = 0 \text{ and } x_2 \geq 1\}$ (see Figure 2). The optimal value is 0.

(e) $f(x_1, x_2) = \max\{x_1, x_2\}$.

Solution: Using the drawing (see Figure 2) it seems that the solution is such that:

$$x_1^\star = x_2^\star = \frac{1}{3}. \tag{6}$$

This can be seen by observing that $x_1 \geq \frac{1}{3}$ or $x_2 \geq \frac{1}{3}$ for all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in the feasible set, with equality at $\begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$. Here, it might be hard to use the first order convexity condition, as the objective function is not differentiable (you can use sub-gradients, but it is beyond the scope of class). The corresponding optimal value is $\frac{1}{3}$.

(f) $f(x_1, x_2) = x_1^2 + 9x_2^2$.

Solution: Using the drawing (see Figure 2) it seems that the solution is such that $x_1^\star = \frac{1}{2}$ and $x_2^\star = \frac{1}{6}$. The corresponding optimal value is $\frac{1}{2}$.

This can be verified using the first order convexity condition:

$$\nabla f \left(\frac{1}{2}, \frac{1}{6} \right)^\top \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} \\ \frac{1}{6} \end{bmatrix} \right) \geq 0, \quad \forall (x_1, x_2) \in \mathcal{X}, \tag{7}$$

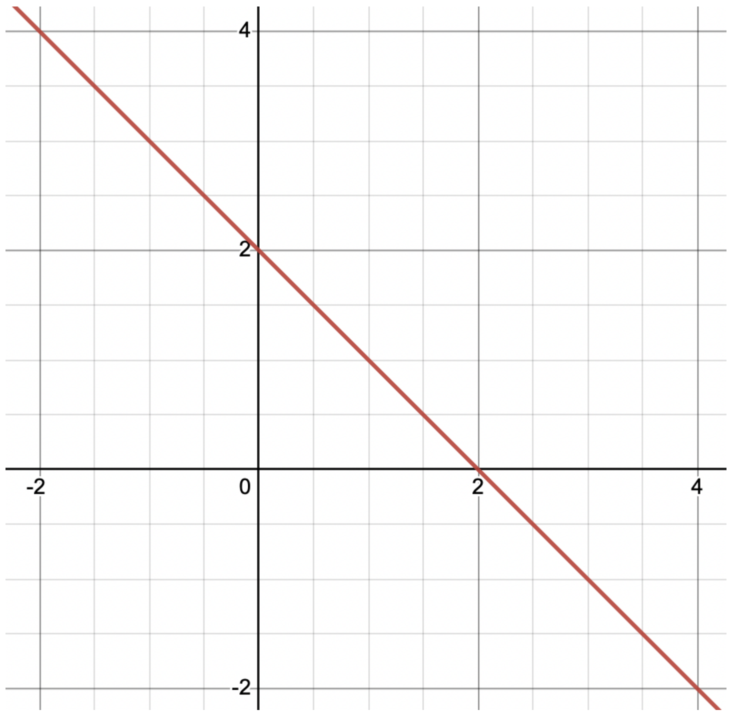
where $\mathcal{X}$ is the feasible set.

2. Fun with Hyperplanes

In this problem we work with hyperplanes, which are key components of linear programming as well as future topics such as support vector machines.

(a) Sketch the affine hyperplane $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^2 \mid \begin{bmatrix} 1 & 1 \end{bmatrix} \vec{x} = 2\}$.

Solution: See the following figure:



(b) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top \vec{x} = 0\}$. Show that $\mathcal{H}$ is a linear subspace of $\mathbb{R}^n$. What is $\dim(\mathcal{H})$?

Solution: We have $\mathcal{H} = \mathcal{N}(\vec{c}^\top)$, where $\vec{c}^\top$ is interpreted as a $1 \times n$ matrix. Thus it is a linear subspace. By the rank-nullity theorem, we have $\dim(\mathcal{R}(\vec{c}^\top)) + \dim(\mathcal{N}(\vec{c}^\top)) = n$, and $\dim(\mathcal{R}(\vec{c}^\top)) = \dim(\mathcal{R}(\vec{c})) = 1$, so $\dim(\mathcal{H}) = \dim(\mathcal{N}(\vec{c}^\top)) = n - 1$.

Alternatively, you could show that $\mathcal{H}$ is closed under linear combinations.

Note that if $\vec{c} \neq \vec{0}$ and $a \neq 0$, then the affine hyperplane $\{\vec{x} \in \mathbb{R}^n | \vec{c}^\top \vec{x} = a\}$ will *not* be a subspace.

(c) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top \vec{x} = 0\}$. Suppose $\vec{x}_\star \in \mathbb{R}^n$ is on one side of the hyperplane, i.e., $\vec{c}^\top \vec{x}_\star > 0$. Give any vector which is on the other side of the hyperplane but not on the hyperplane itself.

Solution: We propose the vector $-\vec{x}_\star$. Indeed, we have

$$\vec{c}^\top (-\vec{x}_\star) = -\vec{c}^\top \vec{x}_\star < 0. \tag{8}$$

Thus $-\vec{x}_\star$ is on the other side of the hyperplane.

(d) Let $\vec{c} \in \mathbb{R}^n$ be nonzero, and let $\vec{x}_0 \in \mathbb{R}^n$ be arbitrary. Let $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top (\vec{x} - \vec{x}_0) = 0\}$. Suppose $\vec{x}_\star \in \mathbb{R}^n$ is on one side of this affine hyperplane. Give any vector which is on the other side of $\mathcal{H}$ but not on $\mathcal{H}$ itself.

Solution: Suppose that without loss of generality we have $\vec{c}^\top (\vec{x}_\star - \vec{x}_0) > 0$. Then we have

$$\vec{c}^\top (\vec{x}_\star - \vec{x}_0) > 0 \implies \vec{c}^\top (\vec{x}_0 - \vec{x}_\star) < 0. \tag{9}$$

We want to find $\vec{z}$ such that $\vec{z} - \vec{x}_0 = \vec{x}_0 - \vec{x}_\star$. By algebra, $\vec{z} = 2\vec{x}_0 - \vec{x}_\star$. This gives $\vec{c}^\top(\vec{z} - \vec{x}_0) = \vec{c}^\top(\vec{x}_0 - \vec{x}_\star) < 0$, so $\vec{z}$ can serve as a vector of the kind that we want.

(e) Let $\vec{x}_0 \in \mathbb{R}^n$ be arbitrary. For a nonzero vector $\vec{c} \in \mathbb{R}^n$, let $\mathcal{H}(\vec{c}) \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top(\vec{x} - \vec{x}_0) = 0\}$. Show that $\vec{0} \in \mathcal{H}(\vec{c})$ for every nonzero $\vec{c} \in \mathbb{R}^n$ if and only if $\vec{x}_0 = \vec{0}$.

Solution: We first claim that, for a fixed nonzero $\vec{c} \in \mathbb{R}^n$, we have $\vec{0} \in \mathcal{H}(\vec{c})$ if and only if $\vec{x}_0$ is orthogonal to $\vec{c}$. Indeed,

$$\begin{aligned}
\vec{c}^\top\vec{x}_0 = 0 &\iff -\vec{c}^\top\vec{x}_0 = 0 &(10)\\
&\iff \vec{c}^\top\vec{0} - \vec{c}^\top\vec{x}_0 = 0 &(11)\\
&\iff \vec{c}^\top(\vec{0} - \vec{x}_0) = 0 &(12)\\
&\iff \vec{0} \in \mathcal{H}(\vec{c}). &(13)
\end{aligned}$$

Thus, if $\vec{c} \neq \vec{0}$, then $\vec{0} \in \mathcal{H}(\vec{c})$ if and only if $\vec{x}_0$ is orthogonal to $\vec{c}$. But if $\vec{x}_0$ is orthogonal to every nonzero $\vec{c} \in \mathbb{R}^n$, then $\vec{x}_0 = \vec{0}$, and the claim is proved.

Note that even for $\vec{c} = \vec{0}$ we can define $\{\vec{x} \in \mathbb{R}^n | \vec{c}^\top(\vec{x} - \vec{x}_0 = 0\}$, but this set will just be $\mathbb{R}^n$, as opposed to the case when $\vec{c} \neq \vec{0}$.

3. Quadratic inequalities

Consider the set S defined by the following inequalities:

$$(x_1 \geq -x_2 + 1 \text{ and } x_1 \leq 0) \text{ or } (x_1 \leq -x_2 + 1 \text{ and } x_1 \geq 0)\,. \tag{14}$$

To be more precise,

$$S_1 = \{\vec{x} \in \mathbb{R}^2 \mid x_1 \geq -x_2 + 1, x_1 \leq 0\}; \tag{15}$$

$$S_2 = \{\vec{x} \in \mathbb{R}^2 \mid x_1 \leq -x_2 + 1, x_1 \geq 0\}; \tag{16}$$

$$S = S_1 \cup S_2. \tag{17}$$

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

Solution:

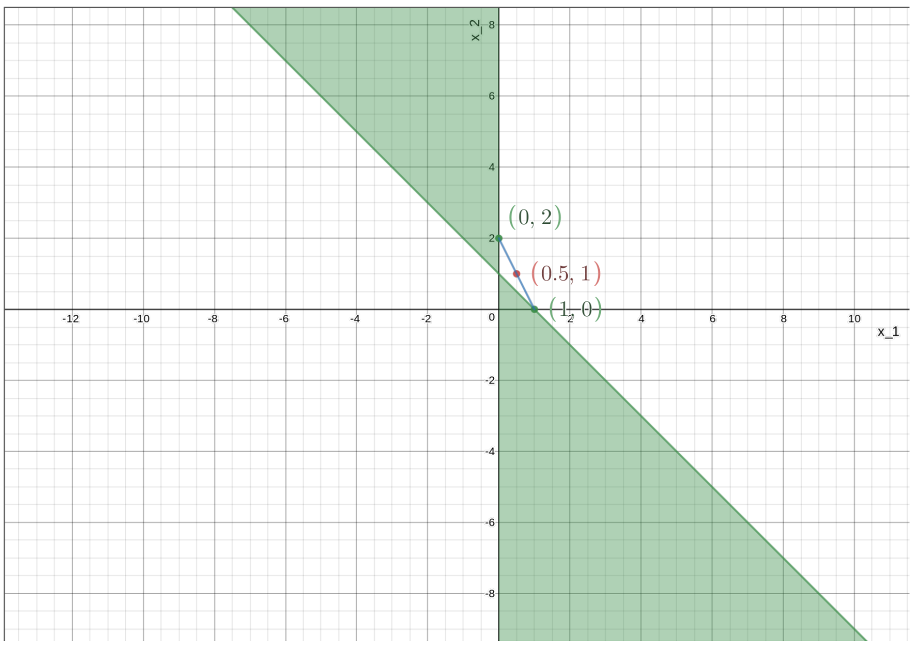


Figure 3: Set S

The set S as shown in Fig. 3 is not convex. We can prove this by providing a 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 via 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 some $A = A^\top \in \mathbb{R}^{2\times 2}$, $\vec{b} \in \mathbb{R}^2$ 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 $q(\vec{x}) := x_1(x_1 + x_2 - 1) \leq 0$ if and only if it is in the set. Expressing $q(\vec{x})$ in the desired form:

$$q(\vec{x}) = x_1^2 + x_1 x_2 - x_1 = \vec{x}^\top A\vec{x} + 2\vec{b}^\top \vec{x} + c$$

where

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

(c) What is the convex hull of 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 an optimization problem. For each of the following optimization problems find the optimal point, $\vec{x}^*$. Describe the constraints that are active in attaining the optimal value. *Hint: Suppose that there exists a point $\vec{x}$ such that $\nabla f(\vec{x}) = 0$. From the first order characterization of a convex function $\vec{x}$ would be an optimum value for f subject to no constraints. If $\vec{x}$ 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 $f(\vec{x}) = (x_1 + 1)^2 + (x_2 - 3)^2$ subject to $\vec{x} \in S_1$.

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 $\vec{x}^* = (-1, 3)$. Now, since $\vec{x}^* \in S_1$, $\vec{x}^*$ is the solution to the constrained optimization problem as well.

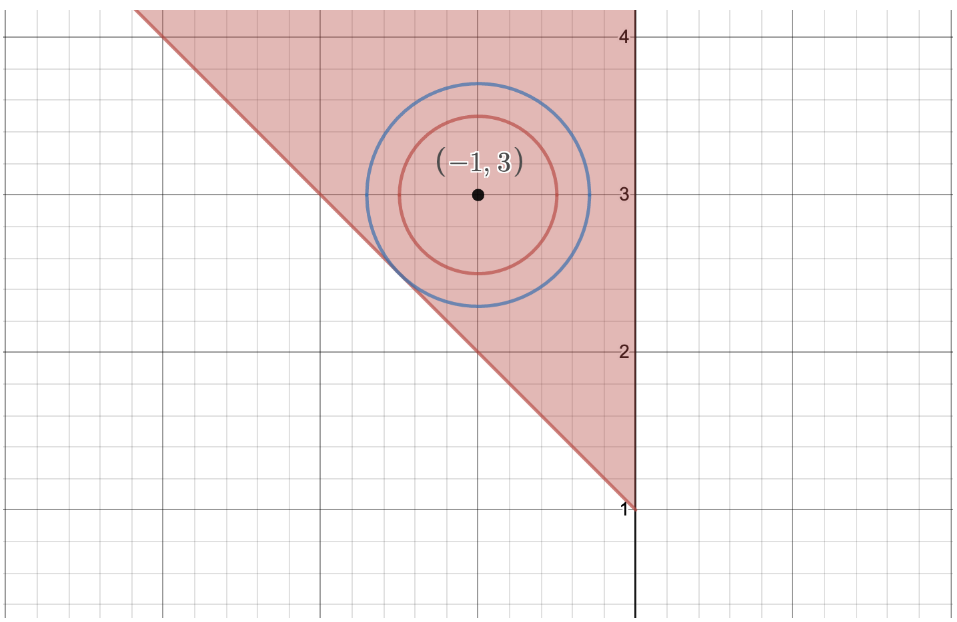


Figure 4: This figure illustrates the position of the optimum, $x^* = (-1, 3)$, and the level sets of the objective function, f, which are concentric circles around x^*.

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 α. 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 $(-3/2, 5/2)$, which is in S_1, 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 $(0, 2)$, which is in S_1, 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 $\vec{x}^*$.

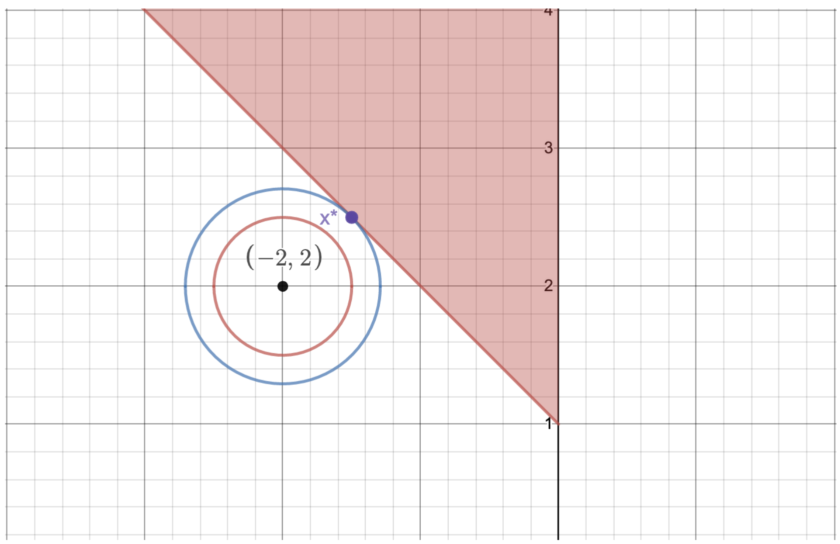


Figure 5: This figure illustrates the position of the optimum, $x^* = (-3/2, 5/2)$, and the level sets of the objective function, f, which are concentric circles around $(-2, 2)$. Note that in this case, the unconstrained optimum does not lie in the set, S_1 and the optimal point lies on the boundary 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 α by setting its gradient with respect to α and setting it to 0, we get the optimal setting of α 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 $\vec{x}^* = (0, 1)$ and both constraints are active at $\vec{x}^*$.

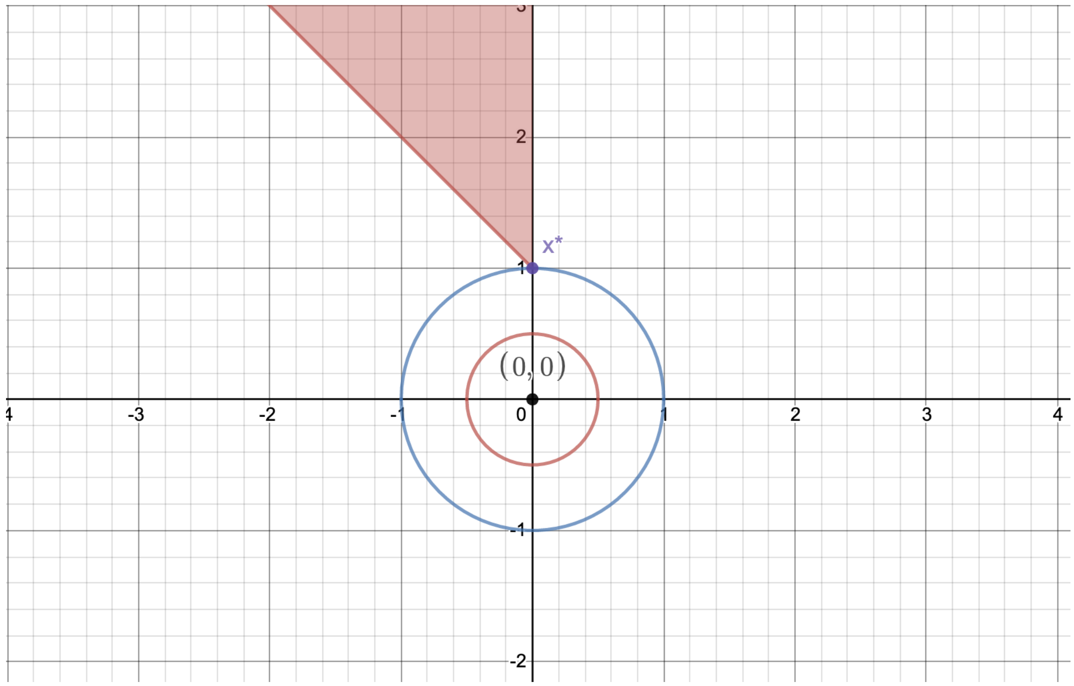


Figure 6: This figure illustrates the position of the optimum, $x^* = (0, 1)$, and the level sets of the objective function, f, which are concentric circles around $(0, 0)$. Note that in this case, the unconstrained optimum does not lie in the set, S_1 and the optimal point lies on the boundary of *both* of the constraints.

4. Gradient Descent Algorithm

Given a continuous and differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient of f at any point $\vec{x}$, $\nabla f(\vec{x})$, is orthogonal to the level curve of f at point $\vec{x}$, and it points in the increasing direction of f. In other words, moving from point $\vec{x}$ in the direction $\nabla f(\vec{x})$ leads to an increase in the value of f, while moving in the direction of $-\nabla f(\vec{x})$ decreases the value of f. This idea leads to an iterative algorithm to minimize the function f: the gradient descent algorithm.

(a) Consider $f(x) = \frac{1}{2}(x-2)^2$, and assume that we use the gradient descent algorithm:

$$x_{k+1} = x_k - \eta \nabla f(x_k) \quad \forall k \geq 0, \tag{18}$$

with some random initialization x_0, where $\eta > 0$ is the step size of the algorithm. Write $(x_k - 2)$ in terms of $(x_0 - 2)$, and show that x_k converges to 2, which is the unique minimizer of f, when $\eta = 0.2$.

Solution: For the given function, we have $\nabla f(x) = (x-2)$; therefore, the gradient descent algorithm gives

$$x_{k+1} = x_k - \eta(x_k - 2). \tag{19}$$

By subtracting 2 from both sides, we obtain

$$(x_{k+1} - 2) = (1-\eta)(x_k - 2) \implies (x_k - 2) = (1-\eta)^k (x_0 - 2). \tag{20}$$

Given $\eta = 0.2$, we have

$$|x_k - 2| = 0.8^k \, |x_0 - 2| \to 0 \quad \text{ as } k \to \infty, \tag{21}$$

which shows that x_k converges to 2.

(b) What is the condition that the step size η must satisfy, to ensure that the gradient descent algorithm converges to 2 from all possible initializations in $\mathbb{R}$? What happens if we choose a larger step size?

Solution: From the solution for part (a), we have

$$|x_k - 2| = |1-\eta|^k \, |x_0 - 2| \quad \forall k \in \mathbb{N}. \tag{22}$$

For convergence of the algorithm for every initialization, it is necessary and sufficient to have $|1-\eta| < 1$, which is equivalent to $\eta \in (0, 2)$. If $\eta = 2$, x_k oscillates around 2 while $|x_k - 2|$ remains fixed. If $\eta > 2$, x_k oscillates around 2 while $|x_k|$ grows unboundedly.

(c) Now assume that we use the gradient descent algorithm to minimize $f(\vec{x}) = \frac{1}{2}\left\|A\vec{x} - \vec{b}\right\|_2^2$ for some $A \in \mathbb{R}^{m \times n}$ and $\vec{b} \in \mathbb{R}^m$, where A has full column rank. First compute $\nabla f(\vec{x})$. Note that $(A^\top A)^{-1}A^\top \vec{b}$ is the solution to the least-squares problem, and $(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b})$ can be thought of as the error at time k relative to the solution. Write $(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b})$ in terms of $(\vec{x}_0 - (A^\top A)^{-1}A^\top b)$.

Solution: We can write $f(\vec{x}) = \frac{1}{2}(\vec{x}^\top A^\top A\vec{x} - \vec{x}^\top A^\top \vec{b} - \vec{b}^\top A\vec{x} + \vec{b}^\top \vec{b})$, so

$$\nabla f(\vec{x}) = A^\top A\vec{x} - A^\top \vec{b}. \tag{23}$$

Then the gradient descent algorithm gives

$$\vec{x}_{k+1} = \vec{x}_k - \eta\left(A^\top A\vec{x}_k - A^\top b\right) = \vec{x}_k - \eta A^\top A\left(\vec{x}_k - (A^\top A)^{-1}A^\top \vec{b}\right). \tag{24}$$

By subtracting $(A^\top A)^{-1}A^\top\vec{b}$ from both sides, we obtain

$$\left(\vec{x}_{k+1} - (A^\top A)^{-1}A^\top\vec{b}\right) = \left(I - \eta A^\top A\right)\left(\vec{x}_k - (A^\top A)^{-1}A^\top\vec{b}\right) \tag{25}$$

and consequently,

$$\left(\vec{x}_k - (A^\top A)^{-1}A^\top\vec{b}\right) = \left(I - \eta A^\top A\right)^k\left(\vec{x}_0 - (A^\top A)^{-1}A^\top\vec{b}\right). \tag{26}$$

(d) Now consider $f(\vec{x}) = \frac{1}{2}\left\|A\vec{x} - \vec{b}\right\|_2^2 + \frac{1}{2}\lambda\left\|\vec{x}\right\|_2^2$ for some $A \in \mathbb{R}^{m\times n}$, $\vec{b} \in \mathbb{R}^m$, and scalar $\lambda > 0$, where A has full column rank. Suppose we solve this problem via gradient descent with step-size $\eta = \frac{1}{\sigma_1^2+\lambda}$, where σ_1 is the maximum singular value of A. Show the gradient descent converges.

Solution: We can write $f(\vec{x}) = \frac{1}{2}(\vec{x}^\top A^\top A\vec{x} - \vec{x}^\top A^\top\vec{b} - \vec{b}^\top A\vec{x} + \vec{b}^\top\vec{b} + \lambda\vec{x}^\top\vec{x})$, so

$$\nabla f(\vec{x}) = A^\top A\vec{x} - A^\top\vec{b} + \lambda\vec{x} = (A^\top A + \lambda I)\vec{x} - A^\top\vec{b}$$

The least-squares solution to this problem is now $\vec{x}^\star = (A^\top A + \lambda I)^{-1}A^\top\vec{b}$. Next, consider the error relative to $\vec{x}^\star$ at time k+1

$$\left(\vec{x}_{k+1} - (A^\top A + \lambda I)^{-1}A^\top\vec{b}\right) = \left(I - \eta(A^\top A + \lambda I)\right)\left(\vec{x}_k - (A^\top A + \lambda I)^{-1}A^\top\vec{b}\right) \tag{27}$$

and consequently,

$$\left(\vec{x}_k - (A^\top A + \lambda I)^{-1}A^\top\vec{b}\right) = \left(I - \eta(A^\top A + \lambda I)\right)^k\left(\vec{x}_0 - (A^\top A + \lambda I)^{-1}A^\top\vec{b}\right). \tag{28}$$

For the algorithm to converge, we need the largest eigenvalue of $(I - \eta(A^\top A + \lambda I))$ to be less than 1 in absolute value. We know any eigenvector $\vec{v}_i$ of $A^\top A$ is also eigenvector of $I - \eta(A^\top A + \lambda I)$ with eigenvalue $1 - \eta(\sigma_i^2 + \lambda)$. With $\eta = \frac{1}{\sigma_1^2+\lambda}$, where σ_1 is the maximum singular value of A given, the following inequality will always be true:

$$\forall i,\ |1 - \eta(\sigma_i^2 + \lambda)| < 1. \tag{29}$$

Therefore, the gradient descent eventually converges to the least-squares solution $\vec{x}^\star$.

5. Gradient Descent Algorithm, continued

Consider $g : \mathbb{R}^n \to \mathbb{R}$, $g(\vec{x}) = \frac{1}{2}\vec{x}^\top M\vec{x} - \vec{x}^\top\vec{b} + c$, where $\vec{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$, and M is a symmetric positive definite matrix, i.e., $M \in \mathbb{S}^n_{++}$.

(a) Write the update rule for the gradient descent algorithm

$$\vec{x}_{k+1} = \vec{x}_k - \eta\nabla g(\vec{x}_k), \tag{30}$$

where η is the step size of the algorithm, and bring it into the form

$$(\vec{x}_{k+1} - \vec{x}_\star) = P_\eta(\vec{x}_k - \vec{x}_\star), \tag{31}$$

where $P_\eta \in \mathbb{R}^{n\times n}$ is a matrix that depends on η. Find $\vec{x}_\star$ and P_η in terms of M, $\vec{b}$, c, and η. *NOTE*: $\vec{x}_\star$ is a minimizer of g.

Solution: We have $\nabla g(\vec{x}) = M\vec{x} - \vec{b}$ and

$$\vec{x}_{k+1} = \vec{x}_k - \eta(M\vec{x}_k - \vec{b}) = \vec{x}_k - \eta M(\vec{x}_k - M^{-1}\vec{b}). \tag{32}$$

We can write

$$\vec{x}_{k+1} - M^{-1}\vec{b} = \vec{x}_k - M^{-1}\vec{b} - \eta M(\vec{x}_k - M^{-1}\vec{b}) = (I - \eta M)(\vec{x}_k - M^{-1}\vec{b}). \tag{33}$$

This shows that $\vec{x}_\star = M^{-1}\vec{b}$ and $P_\eta = I - \eta M$.

(b) Write a condition on the step size η and the matrix M that ensures convergence of $\vec{x}_k$ to $\vec{x}_\star$ for every initialization of $\vec{x}_0$.

Solution: From part (a), we have

$$\vec{x}_k - \vec{x}_\star = (I - \eta M)^k(\vec{x}_0 - \vec{x}_\star). \tag{34}$$

For every initialization $\vec{x}_0$, $(\vec{x}_k - \vec{x}_\star)$ converges to zero if (and only if) all eigenvalues of $(I - \eta M)$ lie in the open interval $(-1, 1)$, i.e., for each eigenvalue λ of M, we require $-1 < 1 - \eta\lambda < 1$.

Since M is positive definite, all of its eigenvalues are positive, and the right hand side of the inequality is satisfied for all $\eta > 0$. For the left hand side of the inequality, we need each eigenvalue λ of M to satisfy $-1 < 1 - n\lambda$. An equivalent condition is:

$$\eta < \frac{2}{\lambda_{\max}(M)}.$$

(c) Assume that all the eigenvalues of M are distinct. Let η_m denote the largest stepsize that ensures convergence for all initializations $\vec{x}_0$, based on the condition computed in part (b).

Does there exist an initialization $\vec{x}_0 \neq \vec{x}_\star$ for which the algorithm converges to the minimum value of g for certain values of the step size η that are larger than η_m?

Justify your answer. *HINT: The question asks if such initializations exist; not whether it is practical to find them.*

Solution: From part (a), we have

$$\vec{x}_k - \vec{x}_\star = (I - \eta M)^k(\vec{x}_0 - \vec{x}_\star). \tag{35}$$

If we want

$$(I - \eta M)^k(\vec{x}_0 - \vec{x}_\star) \to \vec{0} \quad \text{as } k \to \infty \tag{36}$$

for a specific initialization $\vec{x}_0$, the vector $(\vec{x}_0 - \vec{x}_\star)$ must lie in the eigenspaces of $(I - \eta M)$ corresponding to the eigenvalues in the range $(-1, 1)$. This explanation gets full credit.

For example, if $\frac{2}{\lambda_1} < \eta < \frac{2}{\lambda_2}$, where λ_1 and λ_2 are the largest two eigenvalues of M, we have $(I - \eta M)^k(\vec{x}_0 - \vec{x}_\star) \to \vec{0}$ as long as $(\vec{x}_0 - \vec{x}_\star)$ does not have any component in the eigenspace corresponding to the minimum eigenvalue of $(I - \eta M)$.

