1

Problem

Let X be a nonempty convex set in $\mathbb{R}^n$. Let $\alpha : \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{g} : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}$ be convex and let $\mathbf{h}$ be affine, i.e. $\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$. If System 1 below has no solution $\mathbf{x}$, then System 2 has a solution ($\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu}$). The converse holds if $\lambda_0 > 0$.

System $1 \alpha(\mathbf{x}) < 0, \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}$ for some $\mathbf{x} \in X$

System $2 \lambda_0 \alpha(\mathbf{x}) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^{\mathrm{t}} \mathbf{h}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in X$

$(\lambda_0, \boldsymbol{\lambda}) \geq \mathbf{0}, \quad (\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \neq \mathbf{0}$

Hint: Consider the set $\{(z_1, \mathbf{z}_2, \mathbf{z}_3)$: there exists $\mathbf{x} \in X$ such that $\alpha(\mathbf{x}) < z_1, \mathbf{g}(\mathbf{x}) \leq \mathbf{z}_2, \mathbf{h}(\mathbf{x}) = \mathbf{z}_3\}$.

Solution

Let's define the set C as suggested:

$$C = \left\{(z_1, \mathbf{z}_2, \mathbf{z}_3) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^k \mid \exists \mathbf{x} \in X \text{ s.t. } \alpha(\mathbf{x}) < z_1, \mathbf{g}(\mathbf{x}) \leq \mathbf{z}_2, \mathbf{h}(\mathbf{x}) = \mathbf{z}_3\right\}$$

We first show that set C is convex.

Let $(z_1, \mathbf{z}_2, \mathbf{z}_3)$ and $(z_1', \mathbf{z}_2', \mathbf{z}_3')$ be two points in C. By the definition of C, there exist $\mathbf{x}, \mathbf{x}' \in X$ such that:

- $\alpha(\mathbf{x}) < z_1, \mathbf{g}(\mathbf{x}) \leq \mathbf{z}_2, \mathbf{h}(\mathbf{x}) = \mathbf{z}_3$
- $\alpha(\mathbf{x}') < z_1', \mathbf{g}(\mathbf{x}') \leq \mathbf{z}_2', \mathbf{h}(\mathbf{x}') = \mathbf{z}_3'$

Now, consider a convex combination of these two points for any $\theta \in$:

$$(\theta z_1 + (1-\theta) z_1', \theta \mathbf{z}_2 + (1-\theta) \mathbf{z}_2', \theta \mathbf{z}_3 + (1-\theta) \mathbf{z}_3')$$

Since X is a convex set, $\mathbf{x}_\theta = \theta \mathbf{x} + (1-\theta) \mathbf{x}' \in X$.

By the convexity of α and $\mathbf{g}$, and the affinity of $\mathbf{h}$:

- $\alpha(\mathbf{x}_\theta) \leq \theta \alpha(\mathbf{x}) + (1-\theta) \alpha(\mathbf{x}') < \theta z_1 + (1-\theta) z_1'$

- $\mathbf{g}(\mathbf{x}_\theta) \leq \theta \mathbf{g}(\mathbf{x}) + (1-\theta)\mathbf{g}(\mathbf{x}') \leq \theta \mathbf{z}_2 + (1-\theta)\mathbf{z}_2'$
- $\mathbf{h}(\mathbf{x}_\theta) = \mathbf{A}(\theta \mathbf{x} + (1-\theta)\mathbf{x}') - \mathbf{b} = \theta(\mathbf{A}\mathbf{x} - \mathbf{b}) + (1-\theta)(\mathbf{A}'\mathbf{x}' - \mathbf{b}) = \theta \mathbf{z}_3 + (1-\theta)\mathbf{z}_3'$

These inequalities show that the convex combination of the two points is also in C, thus proving that **C is a convex set**.

Then, we show that **System 1 Has No Solution Implies (0, 0, 0) is Not in the Closure of C**

The statement that System 1 has no solution means there is no $\mathbf{x} \in X$ for which $\alpha(\mathbf{x}) < 0$, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$, and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. This is equivalent to saying that the point **(0, 0, 0) is not in the set C**.

More formally, let's define a related set:

$$C_0 = \{(z_1, \mathbf{z}_2, \mathbf{z}_3) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^k \mid \exists \mathbf{x} \in X \text{ s.t. } \alpha(\mathbf{x}) \leq z_1, \mathbf{g}(\mathbf{x}) \leq \mathbf{z}_2, \mathbf{h}(\mathbf{x}) = \mathbf{z}_3\}$$

The set C_0 is also convex. The condition that System 1 has no solution implies that the point $(0, \mathbf{0}, \mathbf{0})$ is not in the set $\{(z_1, \mathbf{z}_2, \mathbf{z}_3) \in C \mid z_1 \leq 0, \mathbf{z}_2 \leq \mathbf{0}, \mathbf{z}_3 = \mathbf{0}\}$.

This implies that $(0, \mathbf{0}, \mathbf{0})$ is not in the closure of C, denoted $\bar{C}$.

Then, we are **Applying the Separating Hyperplane Theorem**

Since C is a nonempty convex set and $(0, \mathbf{0}, \mathbf{0}) \notin \bar{C}$, by the Separating Hyperplane Theorem, there exists a nonzero vector $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^k$ and a scalar β such that:

$$\lambda_0 z_1 + \boldsymbol{\lambda}^t \mathbf{z}_2 + \boldsymbol{\mu}^t \mathbf{z}_3 \geq \beta$$

for all $(z_1, \mathbf{z}_2, \mathbf{z}_3) \in C$, and

$$\lambda_0 \cdot 0 + \boldsymbol{\lambda}^t \mathbf{0} + \boldsymbol{\mu}^t \mathbf{0} < \beta$$

The second inequality implies that $\beta > 0$. Combining these, we get:

$$\lambda_0 z_1 + \boldsymbol{\lambda}^t \mathbf{z}_2 + \boldsymbol{\mu}^t \mathbf{z}_3 > 0$$

for all $(z_1, \mathbf{z}_2, \mathbf{z}_3) \in C$.

Finally we will **Derive the Conditions of System 2**

Let's analyze the properties of the multipliers $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu})$.

- **Non-negativity of λ_0 and $\boldsymbol{\lambda}$**:
  For any $\mathbf{x} \in X$, we can choose $z_1 > \alpha(\mathbf{x})$, $\mathbf{z}_2 \geq \mathbf{g}(\mathbf{x})$, and $\mathbf{z}_3 = \mathbf{h}(\mathbf{x})$. We can make z_1 and the components of $\mathbf{z}_2$ arbitrarily large. If any component of $(\lambda_0, \boldsymbol{\lambda})$ were negative, we could make

the expression $\lambda_0 z_1 + \boldsymbol{\lambda}^t \mathbf{z}_2 + \boldsymbol{\mu}^t \mathbf{z}_3$ arbitrarily negative, which would contradict the separation inequality. Therefore, we must have $(\lambda_0, \boldsymbol{\lambda}) \geq \mathbf{0}$.

- **The Main Inequality**:
  Now, for any $\mathbf{x} \in X$, we can choose a sequence of points $(z_{1,k}, \mathbf{z}_{2,k}, \mathbf{z}_{3,k}) \in C$ such that $z_{1,k} \to \alpha(\mathbf{x})$, $\mathbf{z}_{2,k} \to \mathbf{g}(\mathbf{x})$, and $\mathbf{z}_{3,k} = \mathbf{h}(\mathbf{x})$. From the separation inequality, we have:
  $\lambda_0 z_{1,k} + \boldsymbol{\lambda}^t \mathbf{z}_{2,k} + \boldsymbol{\mu}^t \mathbf{z}_{3,k} \geq 0$
  Taking the limit as $k \to \infty$, we get:
  $\lambda_0 \alpha(\mathbf{x}) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^t \mathbf{h}(\mathbf{x}) \geq 0$
  This inequality holds for all $\mathbf{x} \in X$.
- **Non-zero Multipliers**:
  The Separating Hyperplane Theorem guarantees that the separating vector is non-zero, so $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \neq \mathbf{0}$.

Thus, we have established all the conditions of System 2.

The Converse

Now, let's prove the converse: if System 2 has a solution with $\lambda_0 > 0$, then System 1 has no solution.

Assume there exists a solution $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu})$ to System 2 with $\lambda_0 > 0$. This means:

$$\lambda_0 \alpha(\mathbf{x}) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^t \mathbf{h}(\mathbf{x}) \geq 0$$

for all $\mathbf{x} \in X$, with $(\lambda_0, \boldsymbol{\lambda}) \geq \mathbf{0}$ and $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \neq \mathbf{0}$.

Now, suppose for the sake of contradiction that System 1 has a solution, i.e., there exists an $\mathbf{x}_0 \in X$ such that:

$$\alpha(\mathbf{x}_0) < 0, \quad \mathbf{g}(\mathbf{x}_0) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{x}_0) = \mathbf{0}$$

Let's evaluate the expression from System 2 at this point $\mathbf{x}_0$:

$$\lambda_0 \alpha(\mathbf{x}_0) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}_0) + \boldsymbol{\mu}^t \mathbf{h}(\mathbf{x}_0)$$

- Since $\lambda_0 > 0$ and $\alpha(\mathbf{x}_0) < 0$, we have $\lambda_0 \alpha(\mathbf{x}_0) < 0$.
- Since $\boldsymbol{\lambda} \geq \mathbf{0}$ and $\mathbf{g}(\mathbf{x}_0) \leq \mathbf{0}$, we have $\boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}_0) \leq 0$.
- Since $\mathbf{h}(\mathbf{x}_0) = \mathbf{0}$, we have $\boldsymbol{\mu}^t \mathbf{h}(\mathbf{x}_0) = 0$.

Combining these, we get:

$$\lambda_0 \alpha(\mathbf{x}_0) + \boldsymbol{\lambda}^t \mathbf{g}(\mathbf{x}_0) + \boldsymbol{\mu}^t \mathbf{h}(\mathbf{x}_0) < 0$$

This contradicts the inequality from System 2, which states that the expression must be greater than or equal to zero for all $\mathbf{x} \in X$. Therefore, our assumption that System 1 has a solution must be false.

$\square$

2

Problem

Let $E = \{i : g_i(\mathbf{x}^*) = 0\} = \{1, \ldots, r\}$ and the vectors

$\nabla g_1(\mathbf{x}^*), \ldots, \nabla g_r(\mathbf{x}^*), \nabla h_1(\mathbf{x}^*), \ldots, \nabla h_k(\mathbf{x}^*)$

are linearly independent. Then the system

$\nabla \mathbf{g}_E(\mathbf{x}^*)\,\mathbf{z} < \mathbf{0}, \quad \nabla \mathbf{h}(\mathbf{x}^*)\,\mathbf{z} = \mathbf{0},$

has a solution $\mathbf{z}$ in $\mathbb{R}^n$.

Hint: Use the separation theorem for a point and a convex set.

Solution

Assume there is no vector $\mathbf{z} \in \mathbb{R}^n$ that satisfies the system.

Which means there is no $\mathbf{z}$ such that:

1. $\nabla g_i(\mathbf{x}^*)^t\,\mathbf{z} < 0$ for all $i \in E = \{1, \ldots, r\}$
2. $\nabla h_j(\mathbf{x}^*)^t\,\mathbf{z} = 0$ for all $j = \{1, \ldots, k\}$

Then, we will **Define a Convex Set**

Let's define a set C in the space $\mathbb{R}^{r+k}$ that represents all possible outcomes of the linear transformations on $\mathbf{z}$:

$$C = \left\{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^r \times \mathbb{R}^k : \mathbf{u} = \nabla \mathbf{g}_E(\mathbf{x}^*)\,\mathbf{z}, \mathbf{v} = \nabla \mathbf{h}(\mathbf{x}^*)\,\mathbf{z} \text{ for some } \mathbf{z} \in \mathbb{R}^n\right\}$$

The set C is the range of a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^{r+k}$. Therefore, C is a subspace of $\mathbb{R}^{r+k}$, which means it is a closed and convex set.

Our assumption that the system has no solution means there is no point $(\mathbf{u}, \mathbf{v}) \in C$ such that $\mathbf{u} < \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$.

Let's define another set, D, representing the outcomes we are looking for:

$$D = \left\{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^r \times \mathbb{R}^k : \mathbf{u} < \mathbf{0}, \mathbf{v} = \mathbf{0}\right\}$$

The set D is a convex set (it is the product of the negative orthant, which is convex, and a point). Our assumption is precisely that the sets C and D are disjoint, i.e., $C \cap D = \emptyset$.

Then, we will **Apply the Separating Hyperplane Theorem**

Since C is a closed convex set and D is a convex set, and they are disjoint, by separating hyperplane theorem, we can show the existence of a non-zero vector $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^r \times \mathbb{R}^k$ and a scalar α such that:

1. $\boldsymbol{\lambda}^t \mathbf{u} + \boldsymbol{\mu}^t \mathbf{v} \geq \alpha$ for all $(\mathbf{u}, \mathbf{v}) \in C$
2. $\boldsymbol{\lambda}^t \mathbf{u} + \boldsymbol{\mu}^t \mathbf{v} \leq \alpha$ for all $(\mathbf{u}, \mathbf{v}) \in \bar{D}$ (the closure of D)

Then, we **Analyze the Separation Inequalities**

- **For the set C:**
  Since C is a subspace, if there is any point $(\mathbf{u}_0, \mathbf{v}_0) \in C$ for which $\boldsymbol{\lambda}^t \mathbf{u}_0 + \boldsymbol{\mu}^t \mathbf{v}_0 \neq 0$, then we can scale this point by any positive or negative scalar. This would make the expression $\boldsymbol{\lambda}^t \mathbf{u} + \boldsymbol{\mu}^t \mathbf{v}$ unbounded above and below, which contradicts the inequality $\boldsymbol{\lambda}^t \mathbf{u} + \boldsymbol{\mu}^t \mathbf{v} \geq \alpha$. Therefore, the expression must be constant on C. Since $(\mathbf{0}, \mathbf{0}) \in C$ (by choosing $\mathbf{z} = \mathbf{0}$), this constant value must be 0. Thus, we must have $\alpha = 0$ and:
  $\boldsymbol{\lambda}^t \nabla \mathbf{g}_E(\mathbf{x}^*) \mathbf{z} + \boldsymbol{\mu}^t \nabla \mathbf{h}(\mathbf{x}^*) \mathbf{z} = 0 \quad \text{for all } \mathbf{z} \in \mathbb{R}^n$
  This can be rewritten as:
  $\left(\sum_{i=1}^r \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^k \mu_j \nabla h_j(\mathbf{x}^*)\right)^t \mathbf{z} = 0 \quad \text{for all } \mathbf{z} \in \mathbb{R}^n$
  This implies that the vector itself must be zero:
  $\sum_{i=1}^r \lambda_i \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^k \mu_j \nabla h_j(\mathbf{x}^*) = \mathbf{0}$
- **For the set D:**
  The closure of D is $\bar{D} = \left\{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^r \times \mathbb{R}^k : \mathbf{u} \leq \mathbf{0}, \mathbf{v} = \mathbf{0}\right\}$. The separation inequality becomes:
  $\boldsymbol{\lambda}^t \mathbf{u} \leq 0 \quad \text{for all } \mathbf{u} \leq \mathbf{0}$
  To satisfy this, every component of $\boldsymbol{\lambda}$ must be non-negative. If some $\lambda_i < 0$, we could choose a vector $\mathbf{u}$ with u_i being a large negative number and other components zero, which would make $\boldsymbol{\lambda}^t \mathbf{u} > 0$, a contradiction. Thus, $\boldsymbol{\lambda} \geq \mathbf{0}$.

Then, we **Derive the Contradiction**

From our assumption that the system has no solution, we have concluded that there exist multipliers $\boldsymbol{\lambda} \in \mathbb{R}^r$ and $\boldsymbol{\mu} \in \mathbb{R}^k$ such that:

1. $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ (from the separation theorem)
2. $\boldsymbol{\lambda} \geq \mathbf{0}$
3. $\sum_{i=1}^{r} \lambda_i \nabla g_i\left(\mathbf{x}^*\right) + \sum_{j=1}^{k} \mu_j \nabla h_j\left(\mathbf{x}^*\right) = \mathbf{0}$

Now, could $\boldsymbol{\lambda}$ be the zero vector, $\boldsymbol{\lambda} = \mathbf{0}$? If so, the equation becomes $\sum_{j=1}^{k} \mu_j \nabla h_j\left(\mathbf{x}^*\right) = \mathbf{0}$. By the problem's premise, the set of all gradients (including the ∇h_j's) is linearly independent. This implies that all μ_j must be zero. This would mean $(\boldsymbol{\lambda}, \boldsymbol{\mu}) = (\mathbf{0}, \mathbf{0})$, which contradicts the fact that the separating vector is non-zero.

Therefore, $\boldsymbol{\lambda} \neq \mathbf{0}$. Since we also have $\boldsymbol{\lambda} \geq \mathbf{0}$, this means at least one $\lambda_i > 0$.

The equation $\sum \lambda_i \nabla g_i\left(\mathbf{x}^*\right) + \sum \mu_j \nabla h_j\left(\mathbf{x}^*\right) = \mathbf{0}$ is a linear combination of the gradient vectors. Because $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$, we have found a non-trivial linear combination of these vectors that equals the zero vector. This is the definition of linear dependence.

This **contradicts** the given condition that the vectors $\nabla g_1\left(\mathbf{x}^*\right), \ldots, \nabla g_r\left(\mathbf{x}^*\right), \nabla h_1\left(\mathbf{x}^*\right), \ldots, \nabla h_k\left(\mathbf{x}^*\right)$ are linearly independent.

$\square$

3

Problem

Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{b}$ be an m vector. Then exactly one of the following two systems has a solution:

$$\begin{array}{lll} \text{System 1} & \mathbf{A}\mathbf{x} = \mathbf{b} & \text{for some } \mathbf{x} \in \mathbb{R}^n \\ \text{System 2} & \mathbf{A}^{\mathrm{t}}\mathbf{y} = \mathbf{0}, \mathbf{b}^{\mathrm{t}}\mathbf{y} = 1 & \text{for some } \mathbf{y} \in \mathbb{R}^m \end{array}$$

Hint: Consider the closed convex set $\{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n\}$.

Solution

We will prove it by Farka's Lemma.

First, we **Reformulate System 1**

Any vector $\mathbf{x} \in \mathbb{R}^n$ can be written as the difference of two non-negative vectors. Let $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$, where $\mathbf{x}^+ \geq \mathbf{0}$ and $\mathbf{x}^- \geq \mathbf{0}$.

- Here, $\mathbf{x}_i^+ = \max(x_i, 0)$ and $\mathbf{x}_i^- = \max(-x_i, 0)$.

Substituting this into System 1 gives:
$\mathbf{A}(\mathbf{x}^+ - \mathbf{x}^-) = \mathbf{b}$

We can write this in matrix form by creating a new, larger matrix and a new variable vector. Let:

- $\mathbf{C} = [\mathbf{A}| - \mathbf{A}]$ (an $m \times 2n$ matrix)
- $\mathbf{z} = \begin{bmatrix} \mathbf{x}^+ \\ \mathbf{x}^- \end{bmatrix}$ (a $2n$-vector)

Now, the equation $\mathbf{Ax} = \mathbf{b}$ is equivalent to the system:
$\mathbf{Cz} = \mathbf{b}$ for some $\mathbf{z} \geq \mathbf{0}$.

This is exactly the form of **System II** in Farkas's Lemma

Then, we **Reformulate System 2**

We try to transform $\mathbf{A}^t\mathbf{y} = \mathbf{1}, \mathbf{b}^t\mathbf{y} = 1$ to $\mathbf{C}^t\mathbf{y} \leq 0, \mathbf{b}^t\mathbf{y} > 0$

From the form of $\mathbf{C}$, we have:

$\mathbf{C}^t\mathbf{y} = \mathbf{A}^t\mathbf{y} - \mathbf{A}^t\mathbf{y} = 0$, which satisfies $\mathbf{C}^t\mathbf{y} \leq 0$

Therefore, this is exactly the form of **System II** in Farkas's Lemma

Then, applying the Farkas's Lemma, exactly one of the two systems has a solution.

$\square$

4

Problem

Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{b}$ be an m vector. Then exactly one of the following two systems has a solution:

System 1 $\quad \mathbf{Ax} \leq \mathbf{b}$ for some $\mathbf{x} \in \mathbb{R}^n$

System 2 $\quad \mathbf{A}^{\mathrm{t}}\mathbf{y} = \mathbf{0}, \mathbf{y} \geq \mathbf{0}, \mathbf{b}^{\mathrm{t}}\mathbf{y} < 0$ $\quad$ for some $\mathbf{y} \in \mathbb{R}^m$

Hint: Let $\mathbf{x} = \mathbf{w} - \mathbf{v}, \mathbf{w}, \mathbf{v} \geq \mathbf{0}$. Or consider a new system $\mathbf{Ax} \leq t\mathbf{b}, t > 0$. Or consider the system $\mathbf{Ax} \leq \mathbf{0}, -\mathbf{Ax} \leq \mathbf{0}, -\mathbf{y} \leq \mathbf{0}, -\mathbf{b}^{\mathbf{t}}\mathbf{y} > 0$.

Solution

First, we **Reformulate System I of Farkas's Lemma**

In Farkas's Lemma, $\mathbf{c}^{\mathbf{t}}\mathbf{x} > 0$ is equivalent with $\mathbf{c}^{\mathbf{t}}\mathbf{x} \geq 1$, therefore, we can transform System I into:

$$[\mathbf{D}| - \mathbf{c}]\mathbf{x} \leq [\mathbf{0}| - 1]$$

which satisfies the System 1 in the problem.

Then, we **Reformulate System II of Farkas's Lemma**

With $\mathbf{D}^{\mathbf{t}}\mathbf{w} = \mathbf{c}$ and $\mathbf{w} \geq \mathbf{0}$, let $\mathbf{y} = \begin{bmatrix} \mathbf{w} \\ 1 \end{bmatrix}$, we can get:

$$\mathbf{A}^t\mathbf{y} = \mathbf{c} - \mathbf{c} = 0, \mathbf{b}^t\mathbf{y} = -1 < 0$$

which satisfies the System 2 in the problem.

Then, applying the Farkas's Lemma, exactly one of the two systems has a solution.

$\square$

5

Problem

$$\min f(\mathbf{x}) = (x_1 - 2)^4 + (x_1 - 2x_2)^2.$$

Solution

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix}$$

$$H(\mathbf{x}) = \begin{bmatrix} 12(x_1-2)^2+2 & -4 \\ -4 & 8 \end{bmatrix}$$

Solve $\nabla f(\mathbf{x}) = \mathbf{0}$:

$\mathbf{x}^* = (2,1)$

$$H(2,1) = \begin{bmatrix} 12(2-2)^2+2 & -4 \\ -4 & 8 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix}$$

$\det(H_2) = (2)(8) - (-4)(-4) = 16 - 16 = 0$

Therefore $\mathbf{x}^* = (2,1)$ is the solution.

6

Problem

(Linear Regression) In the linear regression problem n points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ are given in the xy-plane and it is required to "fit" a straight line $y = ax + b$ to these points in such a way that the sum of the squares of the vertical distances of the given points from the line is minimized. That is, a and b are to be chosen so that

$f(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2$

is minimized. The resulting line is called the linear regression line for the given points. Show that the coefficients a and b of the linear regression line are given by

$b = \bar{y} - a\bar{x}, \quad a = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{n(\bar{x})^2 - \sum_{i=1}^{n} x_i^2},$

where

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$

Solution

We calculate the partial derivatives.

$$\frac{\partial f}{\partial a} = 0 \Longrightarrow 2\sum_{i=1}^{n}\left(ax_i^2 + bx_i - x_i y_i\right) = 0$$
$$\sum_{i=1}^{n}\left(ax_i^2 + bx_i - x_i y_i\right) = 0$$

$\Rightarrow a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i = 0$

$$\frac{\partial f}{\partial b} = 0 \implies 2\sum_{i=1}^{n}(ax_i + b - y_i) = 0$$
$$\sum_{i=1}^{n}(ax_i + b - y_i) = 0$$

$\Rightarrow a\sum_{i=1}^{n} x_i + nb - \sum_{i=1}^{n} y_i = 0$

Combine them together, we get:

$b = \bar{y} - a\bar{x}, \quad a = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{n(\bar{x})^2 - \sum_{i=1}^{n} x_i^2},$

□

7

Problem

Maximize $f(\mathbf{x}) = x_1^2 + x_1x_2 + x_2^2$ subject to
$-3x_1 - 2x_2 + 6 \leq 0, \quad -x_1 + x_2 - 3 \leq 0, \quad x_1 - 2 \leq 0.$

1. Sketch the feasible set.
2. Show that a solution exists.
3. Find the solution.

Solution (1)

The feasible set is defined by the intersection of three half-planes. To sketch it, we first draw the boundary lines for each inequality.

1. **Constraint 1:** $-3x_1 - 2x_2 + 6 \leq 0 \implies 3x_1 + 2x_2 \geq 6$
   - The boundary line is $3x_1 + 2x_2 = 6$.
   - If $x_1 = 0$, then $x_2 = 3$. The y-intercept is (0, 3).
   - If $x_2 = 0$, then $x_1 = 2$. The x-intercept is (2, 0).
   - The inequality $3x_1 + 2x_2 \geq 6$ means the feasible region is on or above this line.
2. **Constraint 2:** $-x_1 + x_2 - 3 \leq 0 \implies x_2 \leq x_1 + 3$
   - The boundary line is $x_2 = x_1 + 3$.
   - If $x_1 = 0$, then $x_2 = 3$. The y-intercept is (0, 3).
   - If $x_2 = 0$, then $x_1 = -3$. The x-intercept is (-3, 0).

- The inequality $x_2 \le x_1 + 3$ means the feasible region is on or below this line.

3. **Constraint 3:** $x_1 - 2 \le 0 \implies x_1 \le 2$
   - The boundary line is $x_1 = 2$, which is a vertical line.
   - The inequality $x_1 \le 2$ means the feasible region is on or to the left of this line.

Finding the Vertices:

The feasible region is a polygon. Its vertices are the intersection points of the boundary lines.

- **Vertex A:** Intersection of $3x_1 + 2x_2 = 6$ and $x_2 = x_1 + 3$.
  - Substitute x_2: $3x_1 + 2(x_1 + 3) = 6 \implies 3x_1 + 2x_1 + 6 = 6 \implies 5x_1 = 0 \implies x_1 = 0$.
  - Then $x_2 = 0 + 3 = 3$.
  - **Vertex A = (0, 3)**
- **Vertex B:** Intersection of $3x_1 + 2x_2 = 6$ and $x_1 = 2$.
  - Substitute x_1: $3(2) + 2x_2 = 6 \implies 6 + 2x_2 = 6 \implies 2x_2 = 0 \implies x_2 = 0$.
  - **Vertex B = (2, 0)**
- **Vertex C:** Intersection of $x_2 = x_1 + 3$ and $x_1 = 2$.
  - Substitute x_1: $x_2 = 2 + 3 = 5$.
  - **Vertex C = (2, 5)**

Sketch:

The feasible set is the unbounded region defined by the vertices A(0,3), B(2,0), and C(2,5), extending upwards and to the left.

Solution (2)

The existence of a solution is not guaranteed by the standard Extreme Value Theorem because the feasible set is **unbounded**. We need a more specific argument.

1. **Analyze the objective function:** $f(\mathbf{x}) = x_1^2 + x_1x_2 + x_2^2$.
   This is a quadratic form. Let's analyze its Hessian matrix to determine its convexity.
   - $\frac{\partial f}{\partial x_1} = 2x_1 + x_2$
   - $\frac{\partial f}{\partial x_2} = x_1 + 2x_2$
   - The Hessian matrix is $H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.
2. **Check for convexity:**
   - The first principal minor is $det(H_1) = 2 > 0$.

- The second principal minor is $det(H_2) = (2)(2) - (1)(1) = 3 > 0$.
  Since the principal minors are all positive, the Hessian is positive definite, which means the function $f(\mathbf{x})$ is **strictly convex**.

3. **Apply the correct theorem:** A fundamental result in optimization states that the maximum of a convex function over a closed, convex, polyhedral set (if it exists) must be attained at one of the **extreme points (vertices)** of the set.

Since our feasible set has a finite number of vertices, we can find the maximum by simply evaluating the function at each vertex. The existence of a finite set of candidates guarantees that a maximum among them exists.

Solution (3)

Based on the reasoning above, we only need to test the value of $f(\mathbf{x})$ at the vertices A, B, and C.

- **At Vertex A = (0, 3):**
  $f(0, 3) = (0)^2 + (0)(3) + (3)^2 = 0 + 0 + 9 = 9$
- **At Vertex B = (2, 0):**
  $f(2, 0) = (2)^2 + (2)(0) + (0)^2 = 4 + 0 + 0 = 4$
- **At Vertex C = (2, 5):**
  $f(2, 5) = (2)^2 + (2)(5) + (5)^2 = 4 + 10 + 25 = 39$

Therefore **The maximum value of the function is 39, which occurs at the point $\mathbf{x} = (2, 5)$.**

placeholder

placeholder

placeholder

placeholder

placeholder

placeholder