

Question 1

(a) OLS Solution

We consider the OLS minimization problem

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2.$$

The first-order conditions are obtained by setting the derivatives with respect to β_0 and β_1 to zero:

$$\begin{aligned} \frac{\partial}{\partial \beta_0} : \quad & -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0 \implies \sum_{i=1}^n y_i = n\beta_0 + \beta_1 \sum_{i=1}^n x_i, \\ \frac{\partial}{\partial \beta_1} : \quad & -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i) = 0 \implies \sum_{i=1}^n x_i y_i = \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2. \end{aligned}$$

Defining

$$S_x = \sum_{i=1}^n x_i, \quad S_{xx} = \sum_{i=1}^n x_i^2, \quad S_y = \sum_{i=1}^n y_i, \quad S_{xy} = \sum_{i=1}^n x_i y_i,$$

the normal equations become

$$\begin{cases} n\beta_0 + S_x \beta_1 = S_y, \\ S_x \beta_0 + S_{xx} \beta_1 = S_{xy}. \end{cases}$$

Solving these, we obtain the OLS estimates:

$$\boxed{\hat{\beta}_1 = \frac{nS_{xy} - S_x S_y}{nS_{xx} - S_x^2}, \quad \hat{\beta}_0 = \frac{S_y S_{xx} - S_x S_{xy}}{nS_{xx} - S_x^2}},$$

or, equivalently, in terms of sample means \bar{x} and \bar{y} :

$$\boxed{\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}.$$

(b) Lagrangian Formulation

For the ridge regression problem

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad \text{subject to} \quad \beta_0^2 + \beta_1^2 \leq t,$$

we form the Lagrangian function:

$$\mathcal{L}(\beta_0, \beta_1, \lambda) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 + \lambda (\beta_0^2 + \beta_1^2 - t),$$

with $\lambda \geq 0$.

(c) Optimality under NDCQ

Taking the partial derivatives of \mathcal{L} with respect to β_0 and β_1 and setting them to zero gives:

$$\frac{\partial \mathcal{L}}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) + 2\lambda \beta_0 = 0,$$

$$\frac{\partial \mathcal{L}}{\partial \beta_1} = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i) + 2\lambda \beta_1 = 0.$$

Using the summary statistics, these equations can be rewritten as:

$$S_y - n\beta_0 - S_x\beta_1 = \lambda\beta_0,$$

$$S_{xy} - S_x\beta_0 - S_{xx}\beta_1 = \lambda\beta_1.$$

Rearranging, we obtain the regularized normal equations:

$$(n + \lambda)\beta_0 + S_x\beta_1 = S_y,$$

$$S_x\beta_0 + (S_{xx} + \lambda)\beta_1 = S_{xy}.$$

In matrix form, letting

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

the solution is given by:

$$\boldsymbol{\beta} = (X^T X + \lambda I)^{-1} X^T \mathbf{y},$$

where I is the 2×2 identity matrix. Explicitly, the solutions are:

$$\beta_0 = \frac{S_y(S_{xx} + \lambda) - S_x S_{xy}}{(n + \lambda)(S_{xx} + \lambda) - S_x^2},$$

$$\beta_1 = \frac{(n + \lambda)S_{xy} - S_x S_y}{(n + \lambda)(S_{xx} + \lambda) - S_x^2}.$$

(d) Comparison and Interpretation

OLS Estimates:

$$\hat{\beta}_0 = \frac{S_y S_{xx} - S_x S_{xy}}{n S_{xx} - S_x^2}, \quad \hat{\beta}_1 = \frac{n S_{xy} - S_x S_y}{n S_{xx} - S_x^2}.$$

Ridge Regression Estimates:

$$\beta_0(\lambda) = \frac{S_y(S_{xx} + \lambda) - S_x S_{xy}}{(n + \lambda)(S_{xx} + \lambda) - S_x^2}, \quad \beta_1(\lambda) = \frac{(n + \lambda)S_{xy} - S_x S_y}{(n + \lambda)(S_{xx} + \lambda) - S_x^2}.$$

When $\lambda = 0$, the ridge estimates reduce to the OLS estimates. For $\lambda > 0$, the additional λ in the denominators increases the effective regularization, shrinking the coefficient estimates towards zero. This shrinkage reduces the variance of the estimates, particularly useful in the presence of multicollinearity or limited data, at the cost of introducing some bias.

Question 2

(a)

Since the original problem is to minimize the portfolio variance

$$V(x, y, z) = x^2 + \frac{1}{2}xy + 4y^2 + 2z^2,$$

we convert it into a maximization problem by maximizing the negative of the objective. That is, we solve

$$\max_{x, y, z} - \left(x^2 + \frac{1}{2}xy + 4y^2 + 2z^2 \right).$$

(b)

The equality constraints are:

$$x + y + z = 1,$$

and

$$1 \cdot x + 10 \cdot y + 6 \cdot z = 4.$$

(c)

We form the Lagrangian

$$\mathcal{L}(x, y, z, \lambda, \mu) = x^2 + \frac{1}{2}xy + 4y^2 + 2z^2 - \lambda(x + y + z - 1) - \mu(x + 10y + 6z - 4).$$

The first-order conditions are:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= 2x + \frac{1}{2}y - \lambda - \mu = 0, \\ \frac{\partial \mathcal{L}}{\partial y} &= \frac{1}{2}x + 8y - \lambda - 10\mu = 0, \\ \frac{\partial \mathcal{L}}{\partial z} &= 4z - \lambda - 6\mu = 0.\end{aligned}$$

Together with the constraints

$$x + y + z = 1 \quad \text{and} \quad x + 10y + 6z = 4,$$

solving this system (through elimination and substitution) yields the optimal allocation

$$x = \frac{25}{48}, \quad y = \frac{29}{192}, \quad z = \frac{21}{64}.$$

(d)

To reduce the problem to a univariate unconstrained optimization, express x and z in terms of y . From the budget constraint

$$x = 1 - y - z,$$

and the return constraint

$$x + 10y + 6z = 4,$$

substitute x to obtain

$$1 - y - z + 10y + 6z = 1 + 9y + 5z = 4,$$

which implies

$$5z = 3 - 9y \implies z = \frac{3 - 9y}{5}.$$

Then,

$$x = 1 - y - \frac{3 - 9y}{5} = \frac{2 + 4y}{5}.$$

Substitute these into $V(x, y, z)$ to get

$$V(y) = \frac{288y^2 - 87y + 22}{25}.$$

Differentiating and setting the derivative to zero leads to

$$y = \frac{87}{576} = \frac{29}{192},$$

and substituting back gives

$$x = \frac{25}{48}, \quad z = \frac{21}{64}.$$

(e)

If the required portfolio return increases from 4 to 4.5, the return constraint becomes

$$x + 10y + 6z = 4.5.$$

By sensitivity analysis, the change in the optimal variance is approximated by the Lagrange multiplier μ associated with the return constraint:

$$\Delta V \approx \mu \Delta R.$$

At the optimum, we have

$$\mu = \frac{-x + 5y}{6} = \frac{5}{128}.$$

Thus, for $\Delta R = 0.5$,

$$\Delta V \approx \frac{5}{128} \times 0.5 = \frac{5}{256} \approx 0.0195.$$

(f)

If the mean return of the third stock increases from 6 to 6.5, the return constraint becomes

$$x + 10y + 6.5z = 4.$$

Denote the change in the coefficient for z by $\delta = 0.5$. By the envelope theorem, the sensitivity of the optimal variance with respect to δ is

$$\frac{\partial V^*}{\partial \delta} = -\mu z.$$

At the optimum, with $z = \frac{21}{64}$ and $\mu = \frac{5}{128}$, we have

$$\frac{\partial V^*}{\partial \delta} = -\frac{5}{128} \cdot \frac{21}{64} = -\frac{105}{8192}.$$

Thus, for $\delta = 0.5$, the change in the minimum variance is approximately

$$\Delta V \approx -\frac{105}{8192} \times 0.5 = -\frac{105}{16384} \approx -0.00641.$$

Question 3

(a)

In addition to the budget constraint

$$x + y + z = 1,$$

we now require the portfolio mean return to be at least 4 and no short-selling is allowed. Thus, the inequality constraints are

$$x + 10y + 6z \geq 4,$$

and

$$x \geq 0, \quad y \geq 0, \quad z \geq 0.$$

(b)

Writing the problem in standard form we have:

$$\begin{aligned} & \text{minimize} && V(x, y, z) = x^2 + \frac{1}{2}xy + 4y^2 + 2z^2, \\ & \text{subject to} && x + y + z = 1, \\ & && x + 10y + 6z \geq 4, \\ & && x \geq 0, \quad y \geq 0, \quad z \geq 0. \end{aligned}$$

For the inequality constraint $x + 10y + 6z \geq 4$, we rewrite it in the form

$$4 - x - 10y - 6z \leq 0.$$

Let λ be the Lagrange multiplier for the equality constraint, $\mu \geq 0$ for the inequality $4 - x - 10y - 6z \leq 0$, and $\alpha, \beta, \gamma \geq 0$ for the non-negativity constraints on x , y , and z (written as $-x \leq 0$, $-y \leq 0$, $-z \leq 0$, respectively).

Then the Lagrangian is

$$\mathcal{L}(x, y, z, \lambda, \mu, \alpha, \beta, \gamma) = x^2 + \frac{1}{2}xy + 4y^2 + 2z^2 - \lambda(x + y + z - 1) + \mu(4 - x - 10y - 6z) - \alpha x - \beta y - \gamma z.$$

The first-order (KKT) conditions are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} : \quad & 2x + \frac{1}{2}y - \lambda - \mu - \alpha = 0, \\ \frac{\partial \mathcal{L}}{\partial y} : \quad & \frac{1}{2}x + 8y - \lambda - 10\mu - \beta = 0, \\ \frac{\partial \mathcal{L}}{\partial z} : \quad & 4z - \lambda - 6\mu - \gamma = 0, \end{aligned}$$

together with the primal feasibility conditions

$$\begin{aligned} x + y + z &= 1, \\ 4 - x - 10y - 6z &\leq 0, \\ x, y, z &\geq 0, \end{aligned}$$

the dual feasibility conditions

$$\mu \geq 0, \quad \alpha \geq 0, \quad \beta \geq 0, \quad \gamma \geq 0,$$

and the complementary slackness conditions

$$\mu(4 - x - 10y - 6z) = 0, \quad \alpha x = 0, \quad \beta y = 0, \quad \gamma z = 0.$$

(c) (Bonus)

Under the assumption that a nondegeneracy constraint qualification (NDCQ) holds, the KKT conditions are both necessary and sufficient for optimality.

Notice that the unconstrained optimum from Question 2 is

$$x = \frac{25}{48}, \quad y = \frac{29}{192}, \quad z = \frac{21}{64},$$

which satisfies

$$x + y + z = 1 \quad \text{and} \quad x + 10y + 6z = 4,$$

with $x, y, z > 0$. Therefore, the mean return constraint is binding and the non-negativity constraints are inactive (i.e., $\alpha = \beta = \gamma = 0$). In this case, the KKT conditions reduce to the same system solved in Question 2:

$$\begin{aligned} 2x + \frac{1}{2}y - \lambda - \mu &= 0, \\ \frac{1}{2}x + 8y - \lambda - 10\mu &= 0, \\ 4z - \lambda - 6\mu &= 0, \end{aligned}$$

with

$$x + y + z = 1 \quad \text{and} \quad x + 10y + 6z = 4,$$

and the complementary slackness condition $\mu(4-x-10y-6z) = 0$ implying that $\mu > 0$ (since $4-x-10y-6z = 0$).

Thus, the optimal solution remains

$$\boxed{x = \frac{25}{48}, \quad y = \frac{29}{192}, \quad z = \frac{21}{64}},$$

with the corresponding multiplier μ as determined previously.

Question 4

(a)

We are given the inequalities

$$x + 2y \leq 1, \quad y - 2x \geq -2 \quad (\text{or } y \geq 2x - 2), \quad x + 5 \geq 0 \quad (\text{or } x \geq -5).$$

Calculations of Extreme Points

1. **Intersection of $x + 2y = 1$ and $y = 2x - 2$:**

Substitute $y = 2x - 2$ into $x + 2y = 1$:

$$x + 2(2x - 2) = 1 \implies x + 4x - 4 = 1 \implies 5x = 5 \implies x = 1.$$

Then, $y = 2(1) - 2 = 0$. Thus, one vertex is $(1, 0)$.

2. **Intersection of $x + 2y = 1$ and $x = -5$:**

Substitute $x = -5$ into $x + 2y = 1$:

$$-5 + 2y = 1 \implies 2y = 6 \implies y = 3.$$

Hence, the vertex is $(-5, 3)$.

3. **Intersection of $y = 2x - 2$ and $x = -5$:**

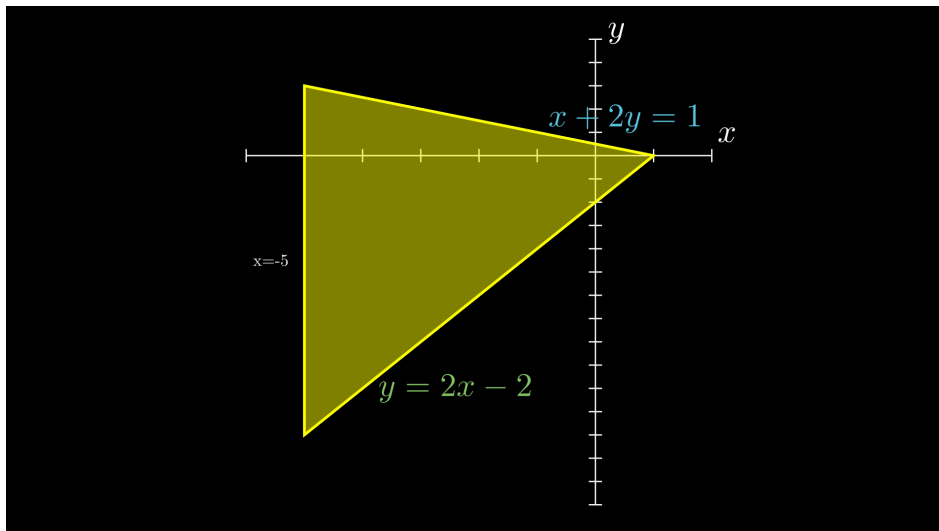
Substitute $x = -5$ into $y = 2x - 2$:

$$y = 2(-5) - 2 = -10 - 2 = -12.$$

Thus, the vertex is $(-5, -12)$.

Summary

- **Sketch:**



The feasible region is the triangle with vertices $(1, 0)$, $(-5, 3)$, and $(-5, -12)$.

- **Convex:** Yes.
- **Extreme Points:** $(1, 0)$, $(-5, 3)$, and $(-5, -12)$.

(b)

We are given the inequalities

$$x + y \geq 0, \quad x - y \leq 0 \quad (\text{or } y \geq x), \quad x \geq 0.$$

Calculations of the Extreme Point

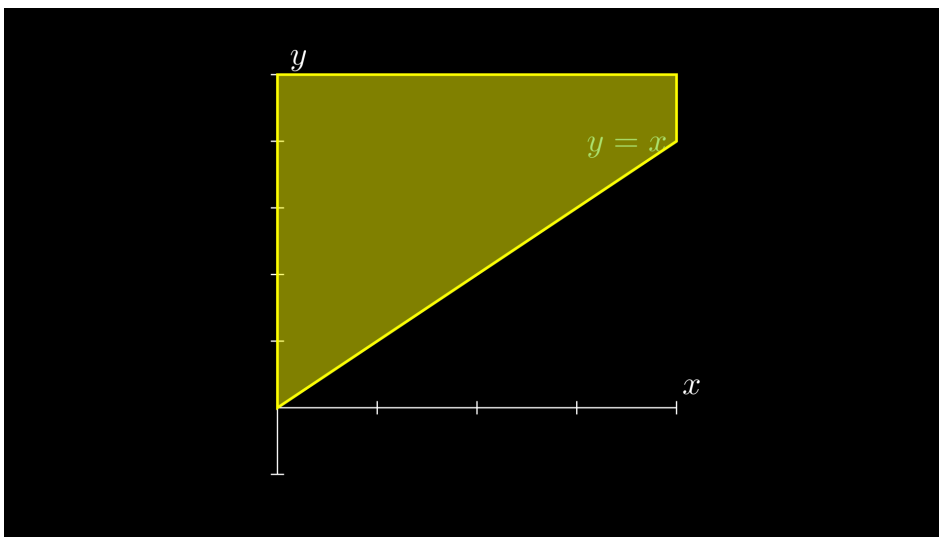
The constraints simplify to $x \geq 0$ and $y \geq x$. Their boundaries intersect when:

$$x = 0 \quad \text{and} \quad y = x,$$

which gives the point $(0, 0)$. This is the only extreme point for this unbounded region.

Summary

- **Sketch:**



The region is the unbounded set defined by $x \geq 0$ and $y \geq x$; that is, the area in the first quadrant above (or on) the line $y = x$.

- **Convex:** Yes.
- **Extreme Point:** $(0, 0)$.

(c) (Bonus)

We are given the inequalities

$$x + y \geq 1, \quad x^2 + y^2 \leq 1.$$

Calculations of Extreme Points

1. Intersection of $x + y = 1$ and $x^2 + y^2 = 1$:

Express $y = 1 - x$ and substitute into $x^2 + y^2 = 1$:

$$x^2 + (1 - x)^2 = 1 \implies x^2 + 1 - 2x + x^2 = 1 \implies 2x^2 - 2x = 0.$$

Factor:

$$2x(x - 1) = 0.$$

Thus, $x = 0$ or $x = 1$. Then,

$$\begin{cases} x = 0 & \Rightarrow & y = 1, \\ x = 1 & \Rightarrow & y = 0. \end{cases}$$

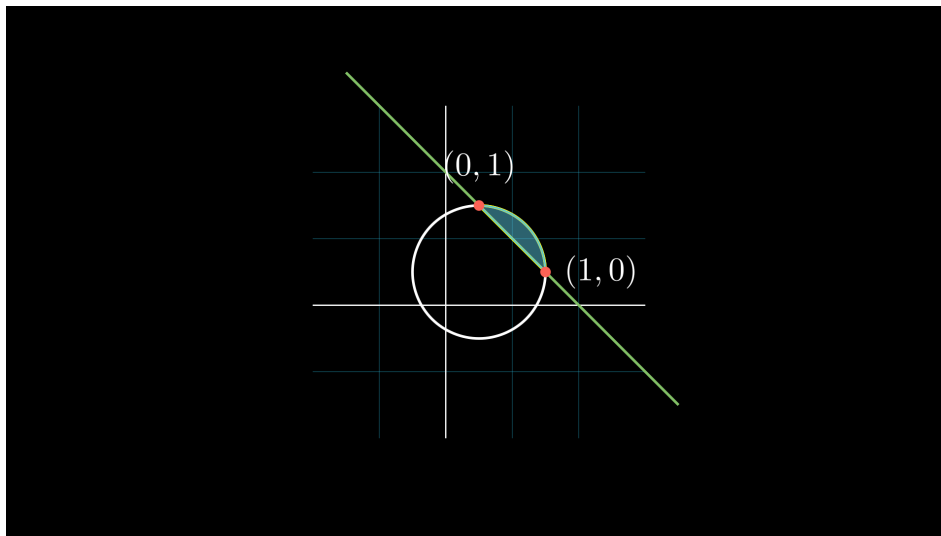
Hence, the intersection points (the endpoints of the chord) are $(0, 1)$ and $(1, 0)$.

2. Other Extreme Points:

Since the unit disk $x^2 + y^2 \leq 1$ is strictly convex, every point on the circular arc that satisfies $x + y \geq 1$ is an extreme point of the intersection. However, for the chord (the line segment between $(0, 1)$ and $(1, 0)$), only the endpoints are extreme.

Summary

- **Sketch:**



The region is the intersection of the unit disk $x^2 + y^2 \leq 1$ and the half-plane $x + y \geq 1$, forming a circular segment (cap) with a boundary consisting of the circular arc from $(0, 1)$ to $(1, 0)$ and the chord joining these points.

- **Convex:** Yes.
- **Extreme Points:** The endpoints $(0, 1)$ and $(1, 0)$ and every point on the circular arc between them.

Question 5

(a) Hyperplane in \mathbb{R}^n

A **hyperplane** in \mathbb{R}^n is defined as the set

$$H = \{x \in \mathbb{R}^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n = b\},$$

where the coefficient vector $a = (a_1, a_2, \dots, a_n)$ is nonzero and b is a real number.

In \mathbb{R}^2 : The equation

$$x + y = 1$$

represents a line (a 1-dimensional hyperplane in \mathbb{R}^2). This line intercepts the x -axis at $(1, 0)$ and the y -axis at $(0, 1)$.

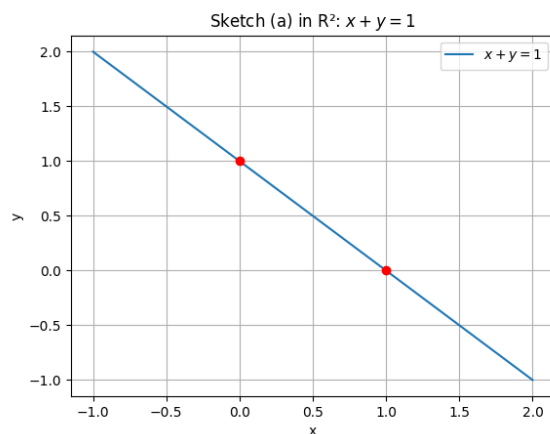


Figure 1: Sketch of $x + y = 1$ in \mathbb{R}^2 .

In \mathbb{R}^3 : The same equation, interpreted in \mathbb{R}^3 , becomes

$$\{(x, y, z) \in \mathbb{R}^3 : x + y = 1\}.$$

This is a plane (a 2-dimensional hyperplane in \mathbb{R}^3) that is “vertical” in the sense that the z -coordinate is free.

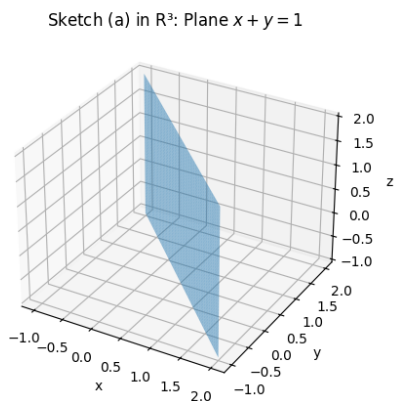


Figure 2: Sketch of the plane $x + y = 1$ in \mathbb{R}^3 .

(b) Half-Spaces Determined by a Hyperplane

Given a hyperplane

$$\{x \in \mathbb{R}^n : a^T x = b\},$$

the two **half-spaces** are defined by the inequalities

$$H^+ = \{x \in \mathbb{R}^n : a^T x \geq b\} \quad \text{and} \quad H^- = \{x \in \mathbb{R}^n : a^T x \leq b\}.$$

For the hyperplane

$$2x + y = 4$$

in \mathbb{R}^2 , the two half-spaces are

$$H^+ = \{(x, y) : 2x + y \geq 4\} \quad \text{and} \quad H^- = \{(x, y) : 2x + y \leq 4\}.$$

Sketch (Half-spaces):

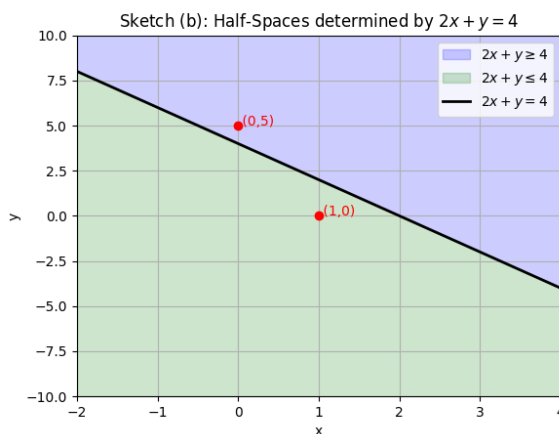


Figure 3: Half-spaces determined by $2x + y = 4$. The point $(1, 0)$ lies in H^- and $(0, 5)$ in H^+ .

For the point $(1, 0)$:

$$2(1) + 0 = 2 < 4 \quad \Rightarrow \quad (1, 0) \in H^-.$$

For the point $(0, 5)$:

$$2(0) + 5 = 5 > 4 \quad \Rightarrow \quad (0, 5) \in H^+.$$

Since these points lie in different half-spaces, the line $2x + y = 4$ **does** separate them.

(c) Hyperplane Separating Two Sets

A hyperplane H is said to **separate** two sets A and B if

$$A \subset H^- \quad \text{and} \quad B \subset H^+ \quad (\text{or vice versa})$$

and neither set crosses H .

The given sets in \mathbb{R}^2 are:

- $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \frac{1}{4}\}$, a closed disk centered at $(0, 0)$ with radius $\frac{1}{2}$.
- $B = \{(x, y) \in \mathbb{R}^2 : (x - 2)^2 + (y - 2)^2 \leq \frac{1}{4}\}$, a closed disk centered at $(2, 2)$ with radius $\frac{1}{2}$.

Checking their positions relative to $2x + y = 4$:

For the center of A , $(0, 0)$:

$$2(0) + 0 = 0 < 4.$$

The distance from $(0, 0)$ to the line is

$$\frac{|0 - 4|}{\sqrt{2^2 + 1^2}} = \frac{4}{\sqrt{5}} \approx 1.788,$$

which is greater than $\frac{1}{2}$. Thus, the entire disk A lies in the half-space $2x + y < 4$.

For the center of B , $(2, 2)$:

$$2(2) + 2 = 6 > 4.$$

The distance from $(2, 2)$ to the line is

$$\frac{|6 - 4|}{\sqrt{5}} = \frac{2}{\sqrt{5}} \approx 0.894,$$

which is also greater than $\frac{1}{2}$. Thus, the entire disk B lies in the half-space $2x + y > 4$.

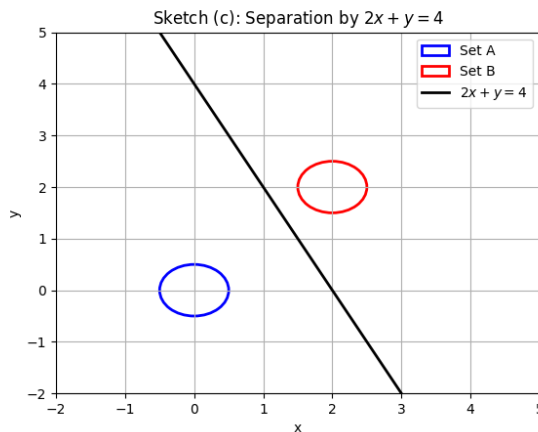


Figure 4: Separation of the disks by the line $2x + y = 4$.

Since neither disk intersects the line and each lies completely in a different half-space, the line $2x + y = 4$ **separates** the two sets.

(d) Separating Hyperplane Theorem and Convexity

The **separating hyperplane theorem** states that if A and B are two nonempty, disjoint, **convex** subsets of \mathbb{R}^n , then there exists a hyperplane that separates them.

Question: Can we remove the condition of “convex sets”?

Answer: No.

Reason: Convexity is crucial. If one (or both) of the sets is nonconvex, even if they are disjoint, there might be no single hyperplane that can separate them.

Counterexample: Consider in \mathbb{R}^2 the following sets:

- $A = \{(x, \sin(1/x)) : x \in (0, 1]\} \cup \{(0, 0)\}$ (nonconvex because of its oscillatory behavior as $x \rightarrow 0^+$).
- $B = \{(x, 0) : x \in [0, 1]\}$, which is convex.

Although A and B are disjoint, the rapid oscillations of A near $x = 0$ prevent any straight line (hyperplane in \mathbb{R}^2) from lying entirely between them.

(e) Supporting and Tangential Hyperplanes

Supporting Hyperplane: A hyperplane H is a **supporting hyperplane** of a set S if

1. H intersects S (i.e., $H \cap S \neq \emptyset$), and
2. S lies entirely in one of the closed half-spaces determined by H .

Tangential Hyperplane: A **tangential hyperplane** is a supporting hyperplane that “just touches” the set S at a boundary point (or along a boundary curve) without cutting through its interior. In classical calculus, a tangent line to a smooth curve is a tangential hyperplane in \mathbb{R}^2 .

Examples in \mathbb{R}^2 :

1. Tangential Hyperplane Example: Let

$$S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \quad (\text{the unit disk}).$$

The line

$$x = 1$$

touches S only at the point $(1, 0)$ and S lies entirely in $\{x \leq 1\}$. Hence, $x = 1$ is a tangential hyperplane.

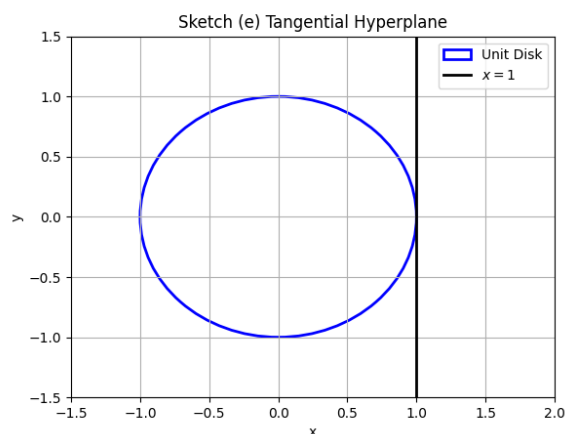


Figure 5: Tangential hyperplane: $x = 1$ touching the unit disk.

2. Supporting (but not Tangential) Hyperplane Example: Consider the square

$$S = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

The line

$$x = 1$$

is a supporting hyperplane since $S \subset \{x \leq 1\}$ and the hyperplane meets S along the entire vertical edge $\{1\} \times [0, 1]$.

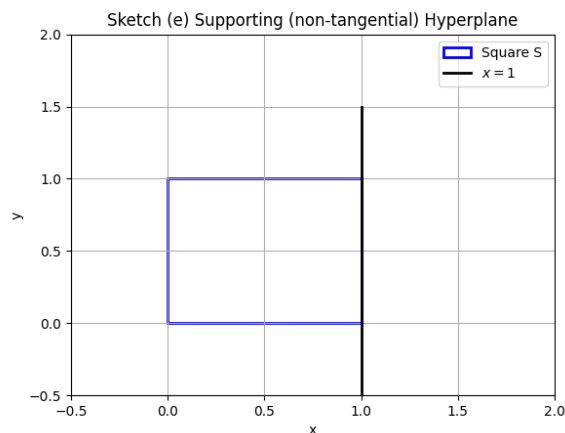


Figure 6: Supporting (non-tangential) hyperplane: $x = 1$ supporting the square.

(f) Extreme Points of a Set

An **extreme point** of a convex set S is a point $x \in S$ that cannot be written as a convex combination

$$x = ty + (1 - t)z, \quad 0 < t < 1,$$

of two distinct points $y, z \in S$.

Consider the set

$$S = \{(x, y) \in \mathbb{R}^2 : x \in [0, \pi], y \leq \sin(x)\}.$$

This is the region in \mathbb{R}^2 that lies below (or on) the curve $y = \sin(x)$ for $x \in [0, \pi]$.

The **upper boundary** of S is given by the graph

$$\{(x, \sin(x)) : x \in [0, \pi]\}.$$

Since $\sin(x)$ is strictly concave on $(0, \pi)$ (because $\sin''(x) = -\sin(x) < 0$ for $x \in (0, \pi)$), every point on the curve is exposed to the top and cannot be expressed as a convex combination of other points in S . In particular, the endpoints $(0, 0)$ and $(\pi, 0)$ are extreme.

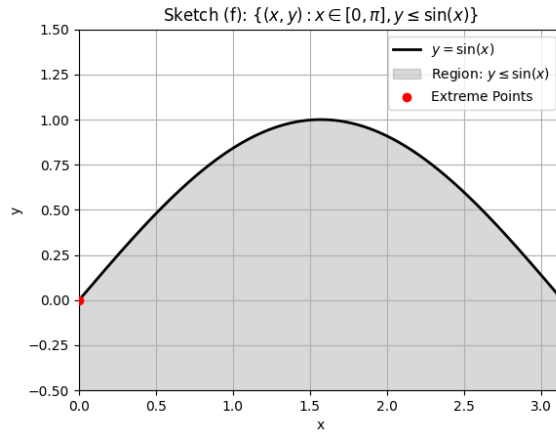


Figure 7: Sketch of the set $S = \{(x, y) : x \in [0, \pi], y \leq \sin(x)\}$, showing the curve $y = \sin(x)$ and extreme points at $(0, 0)$ and $(\pi, 0)$.