Linear transformations

7/9/2020

Francesco Preta

Computational Linear Algebra

Definition

Let V, W be vector spaces. A function $T : V \to W$ is called **linear** if for every $\mathbf{x}, \mathbf{y} \in V$ and $\alpha, \beta \in \mathbb{R}$, we have

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$$

Examples:

Matrix multiplication

For a fixed matrix $A \in \mathbb{R}^{n \times m}$: $A(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha A \mathbf{x} + \beta A \mathbf{y}$

For instance $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$

Examples:

Inner product with a fixed vector

Fix a vector $\mathbf{z} \in \mathbb{R}^n$, then for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$and $\alpha, \beta \in \mathbb{R}$:

$$\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$$

For instance, let $\mathbf{z} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

Examples:

Integral in function space

let $f_1, f_2 : [0,1] \to \mathbb{R}$ be Riemann-integrable functions and $\alpha, \beta \in \mathbb{R}$.
Then

$$\int_0^1 (\alpha f_1(t) + \beta f_2(t))\, dt = \alpha \int_0^1 f_1(t)\, dt + \beta \int_0^1 f_2(t)\, dt$$

Examples:

Derivatives on function spaces

let $g_1, g_2 : [0,1] \to \mathbb{R}$ be derivable functions and $\alpha, \beta \in \mathbb{R}$. Then

$$\frac{d}{dt}(\alpha g_1 + \beta g_2)(t) = \alpha \frac{dg_1(t)}{dt} + \beta \frac{dg_2(t)}{dt}$$

Non-Examples:

Euclidean norm

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ in general it is not true that

$$||\alpha\mathbf{x} + \beta\mathbf{y}|| = \alpha||\mathbf{x}|| + \beta||\mathbf{y}||$$

For instance, consider $\mathbf{x} = [1 \quad 0]$, $\mathbf{y} = [0 \quad 1]$ and $\alpha = \beta = 1$

Non-examples:

Maximum function

Consider $\max : \mathbb{R}^n \to \mathbb{R}$, the function returning the maximum element in a vector. Then for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$and $\alpha, \beta \in \mathbb{R}$, in general

$$\max(\alpha\mathbf{x} + \beta\mathbf{y}) \neq \alpha \max(\mathbf{x}) + \beta \max(\mathbf{y})$$

As a counterexample, consider $\mathbf{x} = [0 \quad 1]$ and $\mathbf{y} = [0 \quad -1]$, $\alpha = \beta = 1$.

Canonical basis of $\mathbb{R}^n$

Finite set of vectors $\{\mathbf{e}_i\}_{i=1}^m$, where

$$(\mathbf{e}_i)_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Then every vector $\mathbf{x} \in \mathbb{R}^m$ is a linear combination of $\mathbf{e}_i$'s, that is

$\mathbf{x} = \sum_{i=1}^m x_i \mathbf{e}_i$.

Representation of linear functions

A priori, it seems to be the case that matrix multiplication constitutes only one of the possible linear functions between $\mathbb{R}^m$ and $\mathbb{R}^n$. However we will proving that every linear function on $\mathbb{R}^m$ is in some sense a matrix multiplication:

Representation Theorem for linear functions

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a linear operator. Then there exists a $n \times m$ matrix A such that $f(\mathbf{x}) = A\mathbf{x}$ for every$\mathbf{x} \in \mathbb{R}^m$

Proof:

Linearity of f guarantees that the value of $f(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^m$ is uniquely determined by the value of f on a finite set of vectors that span the entire space: the canonical basis.

$$f(\mathbf{x}) = f(\sum_{i=1}^{m} x_i \mathbf{e}_i) = \sum_{i=1}^{m} x_i f(\mathbf{e}_i)$$

Proof (continued):

For any matrix $A \in \mathbb{R}^{n \times m}$, $A\mathbf{e}_i$ is the i-th column of A:

$$(A\mathbf{e}_i)_j = \sum_{k=1}^{n} A_{j,k}(\mathbf{e}_i)_k = A_{j,i}$$

Proof (continued):

Then consider the matrix

$$A = \begin{bmatrix} f(\mathbf{e}_1) & f(\mathbf{e}_2) & \dots & f(\mathbf{e}_m) \end{bmatrix}$$

Where every $f(\mathbf{e}_i)$ is a column matrix. We have that

$$A\mathbf{x} = A(\sum_{i=1}^{m} x_i\mathbf{e}_i) = \sum_{i=1}^{m} x_i A\mathbf{e}_i = \sum_{i=1}^{m} x_i f(\mathbf{e}_i) = f(\sum_{i=1}^{m} x_i\mathbf{e}_i) = f(\mathbf{x})$$

Example:

Shift-forward operator

Consider $f : \mathbb{R}^3 \to \mathbb{R}^3$ to be the shift-forward operator, that is

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix}$$

What is the representation of f?

Injectivity and Surjectivity

A function $f : V \to W$ is said to be **injective** if

$$f(\mathbf{x}) = f(\mathbf{y}) \quad \text{if and only if } \mathbf{x} = \mathbf{y}$$

that is, different points of V have different image in W. f is said to be **surjective** if

$$\forall \mathbf{w} \in W, \exists \mathbf{x} \in V \quad \text{such that } f(\mathbf{x}) = \mathbf{w}$$

that is, all points in W are the image of a point in V.

If a function is both injective and surjective it is said to be **bijective**.

Injectivity for linear functions

Linearity guarantees an additional property for surjectivity, that is:

Theorem

Let $T : V \to W$ be a linear function. Then injectivity of T is equivalent to the condition that $T(\mathbf{x}) = 0$ if and only if $\mathbf{x} = 0$.

Proof:

Consider $\mathbf{x}, \mathbf{y} \in V$ such that $T(\mathbf{x}) = T(\mathbf{y})$. Since T is linear and $T(\mathbf{x}) - T(\mathbf{y}) = 0$, then$T(\mathbf{x} - \mathbf{y}) = 0$.

If T is injective, then $\mathbf{x} = \mathbf{y}$, therefore $\mathbf{x} - \mathbf{y} = 0$.

On the other hand, if $T(v) = 0$ if and only if $v = 0$, then $\mathbf{x} - \mathbf{y} = 0$, so $\mathbf{x} = \mathbf{y}$ and T is injective.

The image of 0

for a linear operator it is always the case that $T(\mathbf{0}) = \mathbf{0}$. This comes from the definition of linearity, since

$T(0) = T(2 \cdot 0) = 2T(0)$

which is true only if $T(0) = 0$.

The theorem tells us that injectivity of T corresponds to the fact that $0$is the only element whose image is $\mathbf{0}$.

Kernel and Image

For linear functionals, it makes sense to give the following definitions:

Let $T : V \to W$ be a linear function. Then the **Kernel** of T is defined as the elements of V whose image is $\mathbf{0}$, that is

Ker$(T) = \{\mathbf{x} \in V \,|\, T(\mathbf{x}) = 0\}$

and the **image** or **range** of T are the elements of W that are mapped through T by some element in V

Im$(T) = \{y \in W \,|\, T(\mathbf{x}) = y$ for some $\mathbf{x} \in V\}$

Surjectivity of a linear function

Surjectivity for linear functions can be stated in the following terms:

Theorem

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear operator and A a $n \times m$ matrix such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^m$. Then T is surjective if and only if the columns of A span $\mathbb{R}^n$.

Proof:

By the representation theorem $T(\mathbf{x}) = A\mathbf{x}$ for some matrix A in $\mathbb{R}^{n \times m}$, then every $\mathbf{x}$ can be written as a linear combination of the canonical basis of $\mathbb{R}^m$, and by linearity

$$A\mathbf{x} = A(\sum_{i=1}^{m} x_i \mathbf{e}_i) = \sum_{i=1}^{m} x_i A \mathbf{e}_i$$

But we know that $A\mathbf{e}_i$ is the i-th column of A, so $\mathrm{Im}(T)$ is the span of the columns of A.

Example

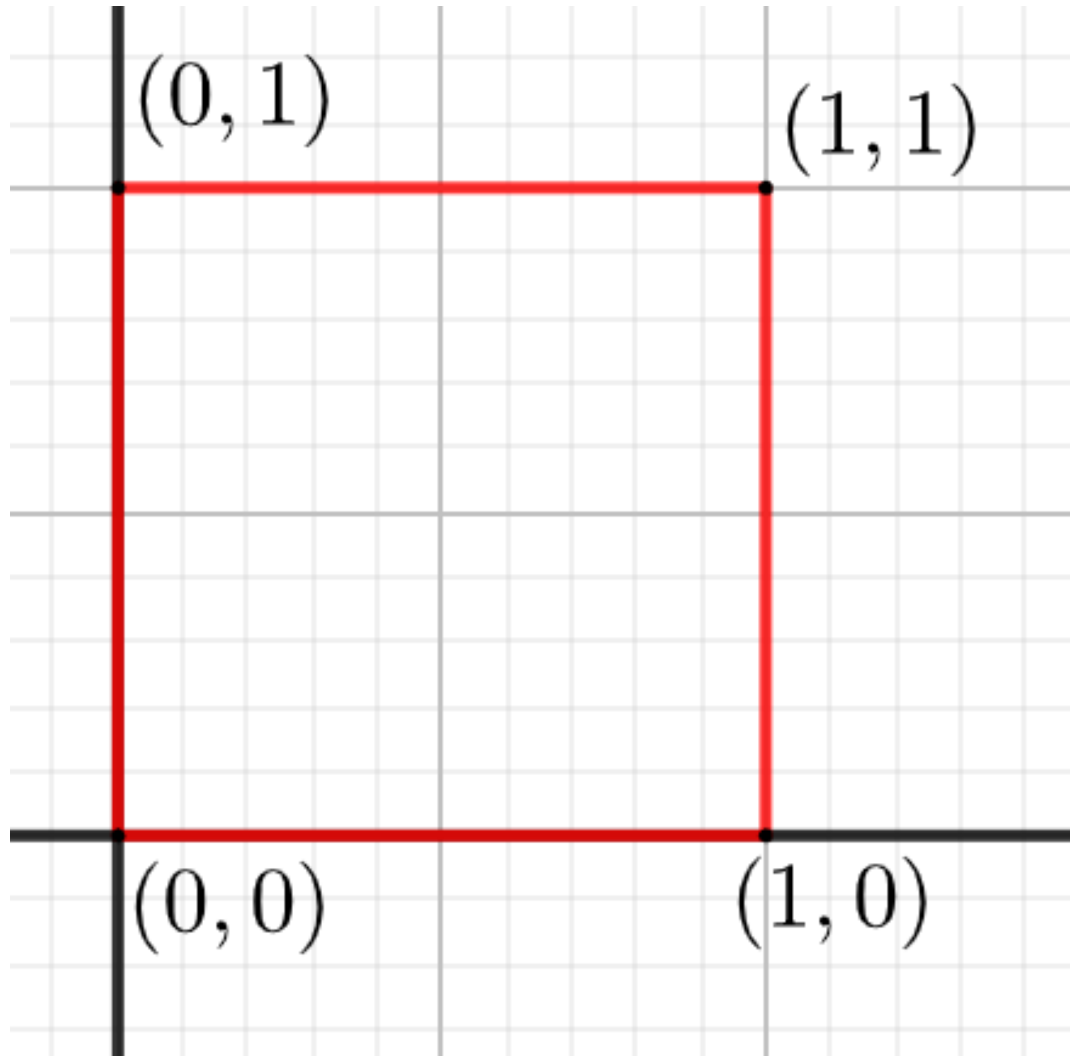
Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}^3$:

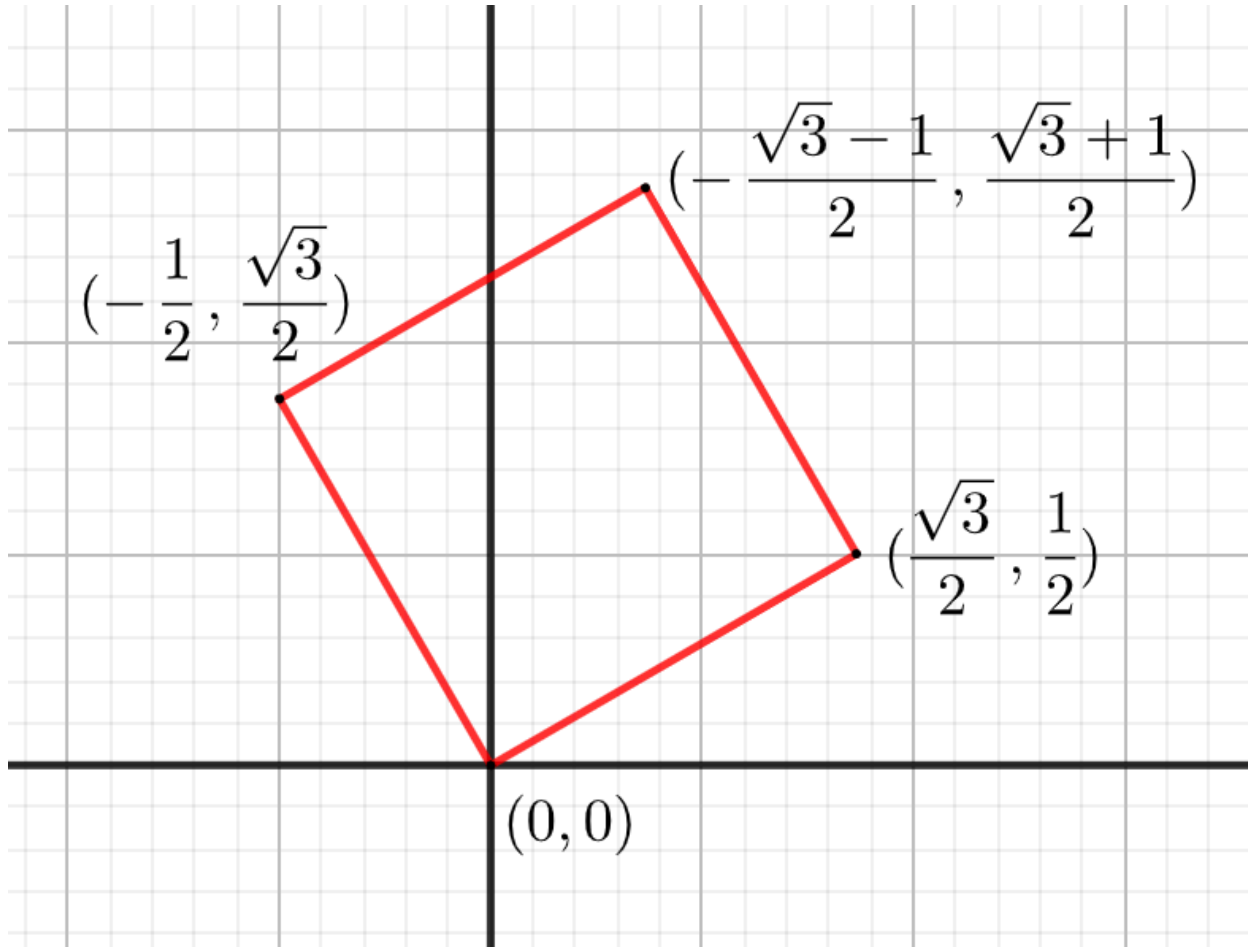
$f(\mathbf{x}) = (x_1, x_1, x_2)$

Is it injective? Is it surjective?

Examples of linear transformations in $\mathbb{R}^2$

Rotations





Example of counter-clockwise rotation by $\frac{\pi}{3}$, for A= $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$.

Rotations

General form of rotation is A= $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

represents a rotation around the origin of angle θ. Consider a point $\mathbf{x} = (x_1, x_2)$. Then

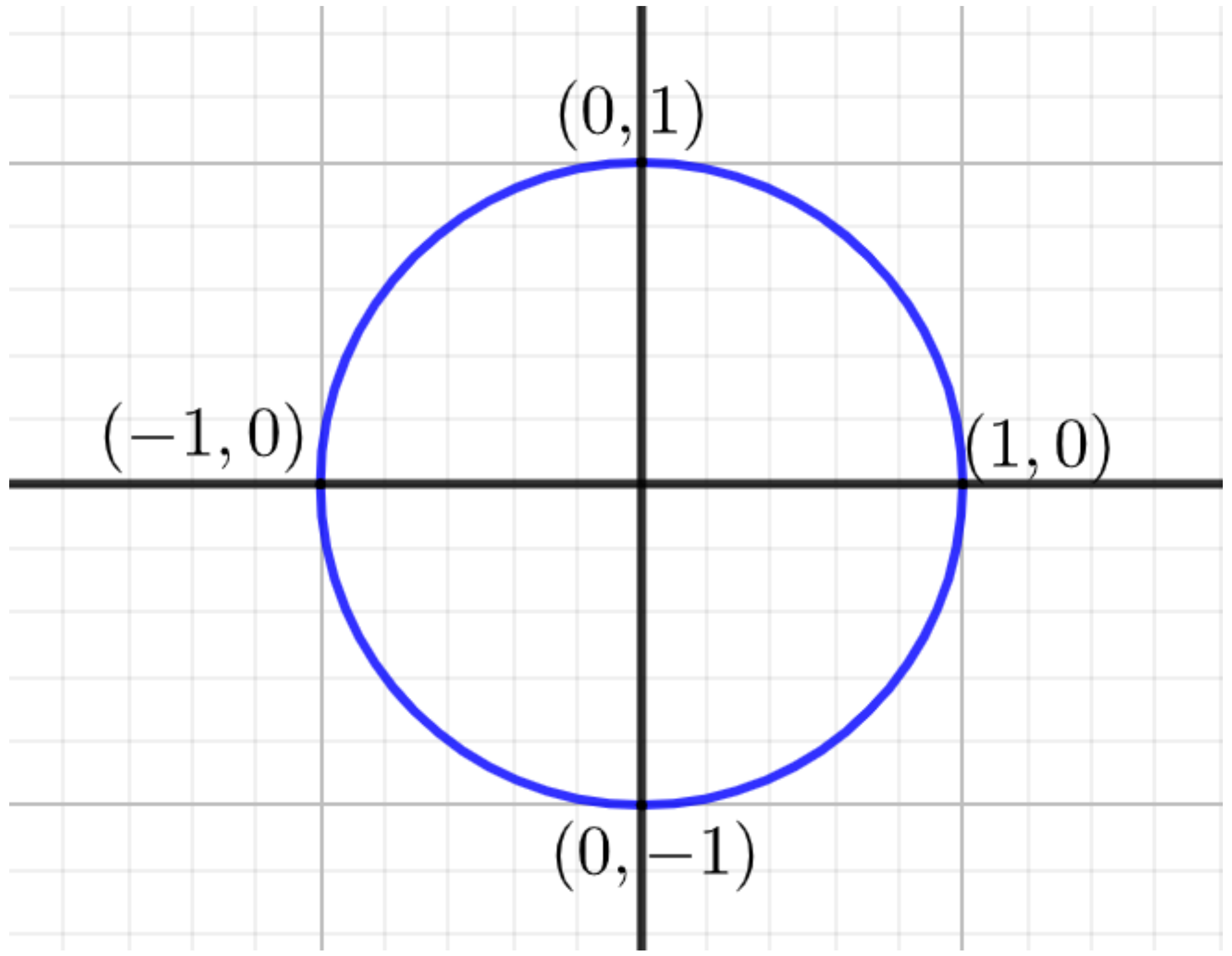
$A\mathbf{x} = (x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta)$

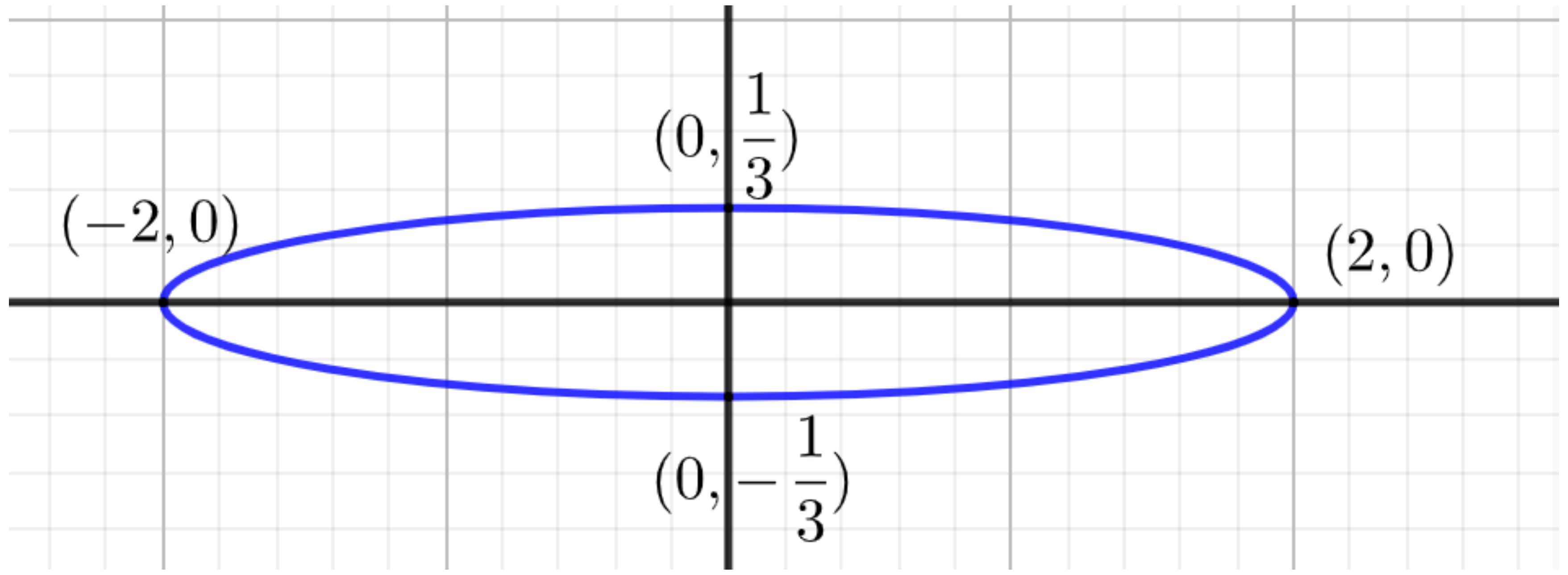
Rotations

To check that it does what we want it to do, we can calculate

$$\cos\angle(\mathbf{x}, A\mathbf{x}) = \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{||\mathbf{x}||||A\mathbf{x}||}$$

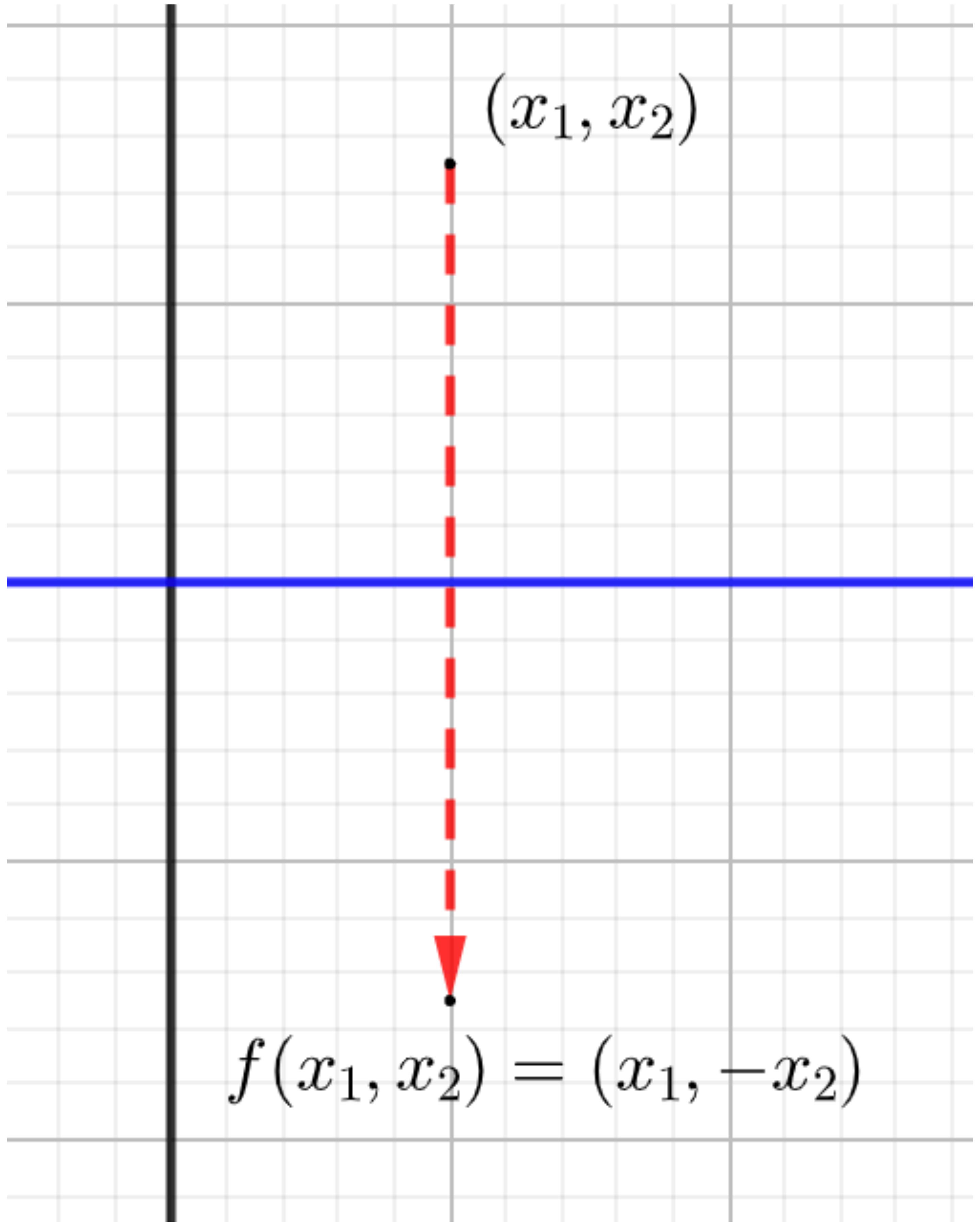
Dilatations





Example of dilatation. General form is $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. In this case the horizontal axis is dilated by 2 and the vertical axis contracted by $\frac{1}{3}$, so the form is $A = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$

Reflections: horizontal axis



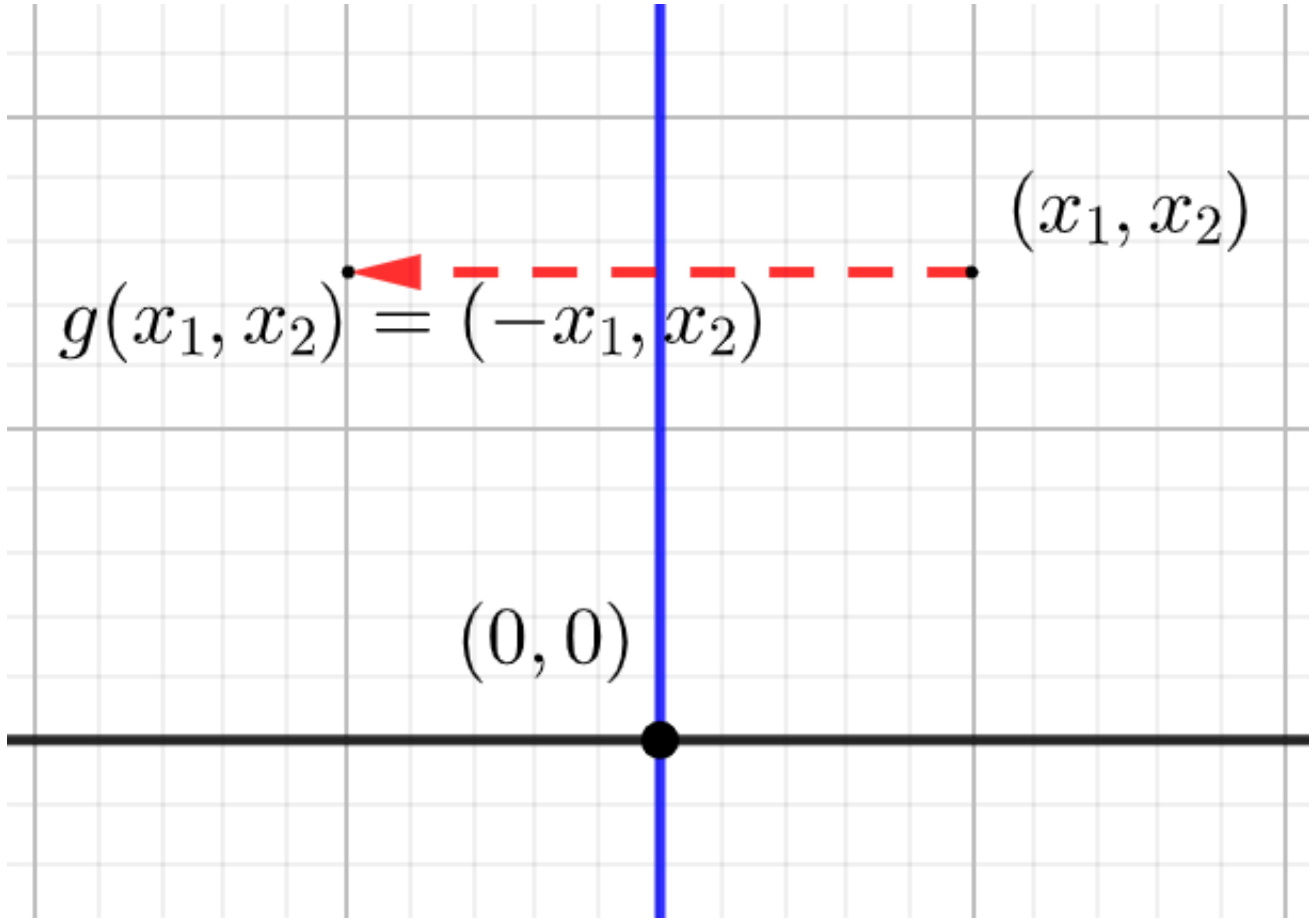
The function has the form

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

The matrix A has the form

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Reflection: vertical axis



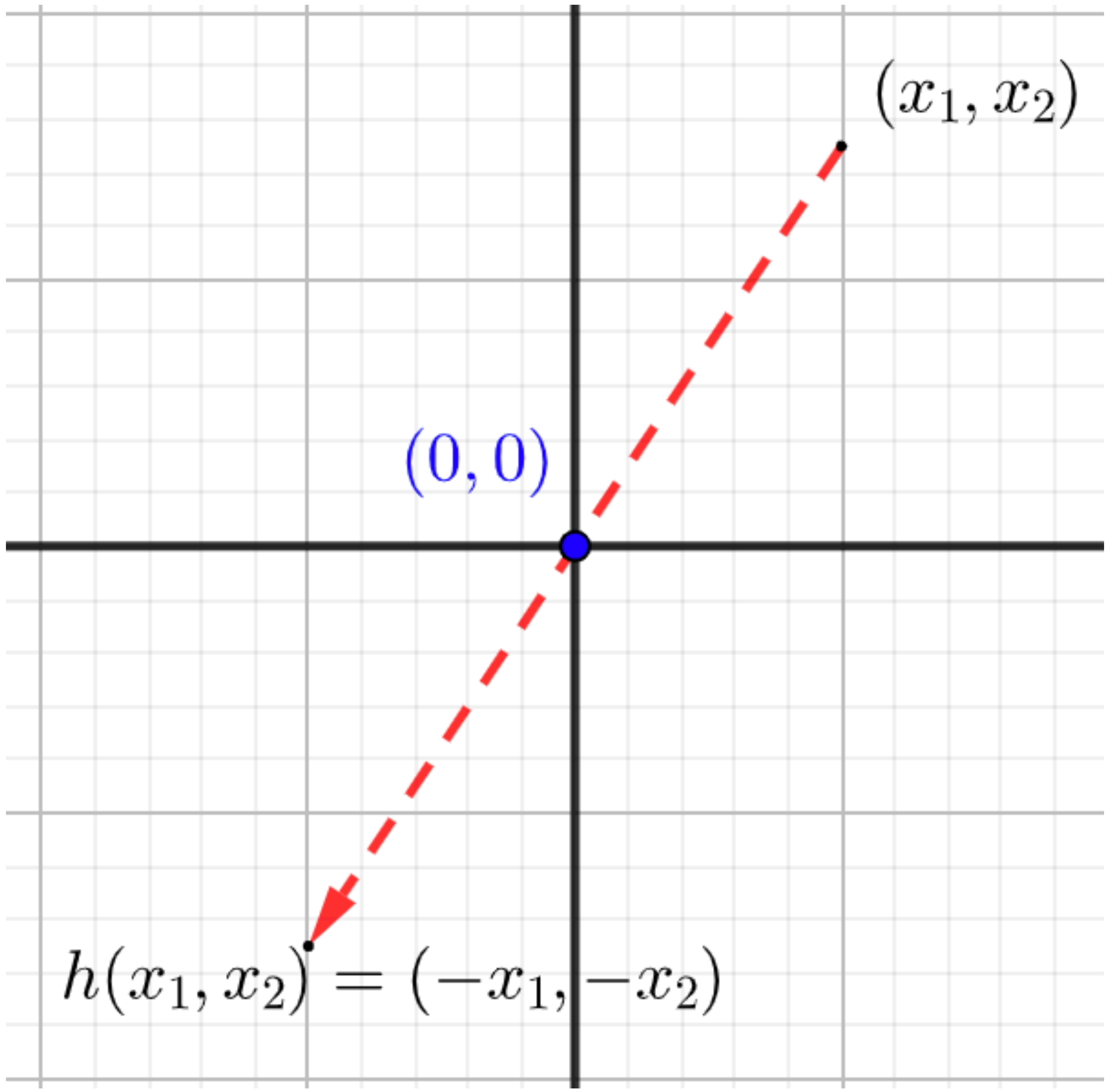
The function has the form

$$g(x_1, x_2) = (-x_1, x_2).$$

The matrix B has the form

$$B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Reflection: origin



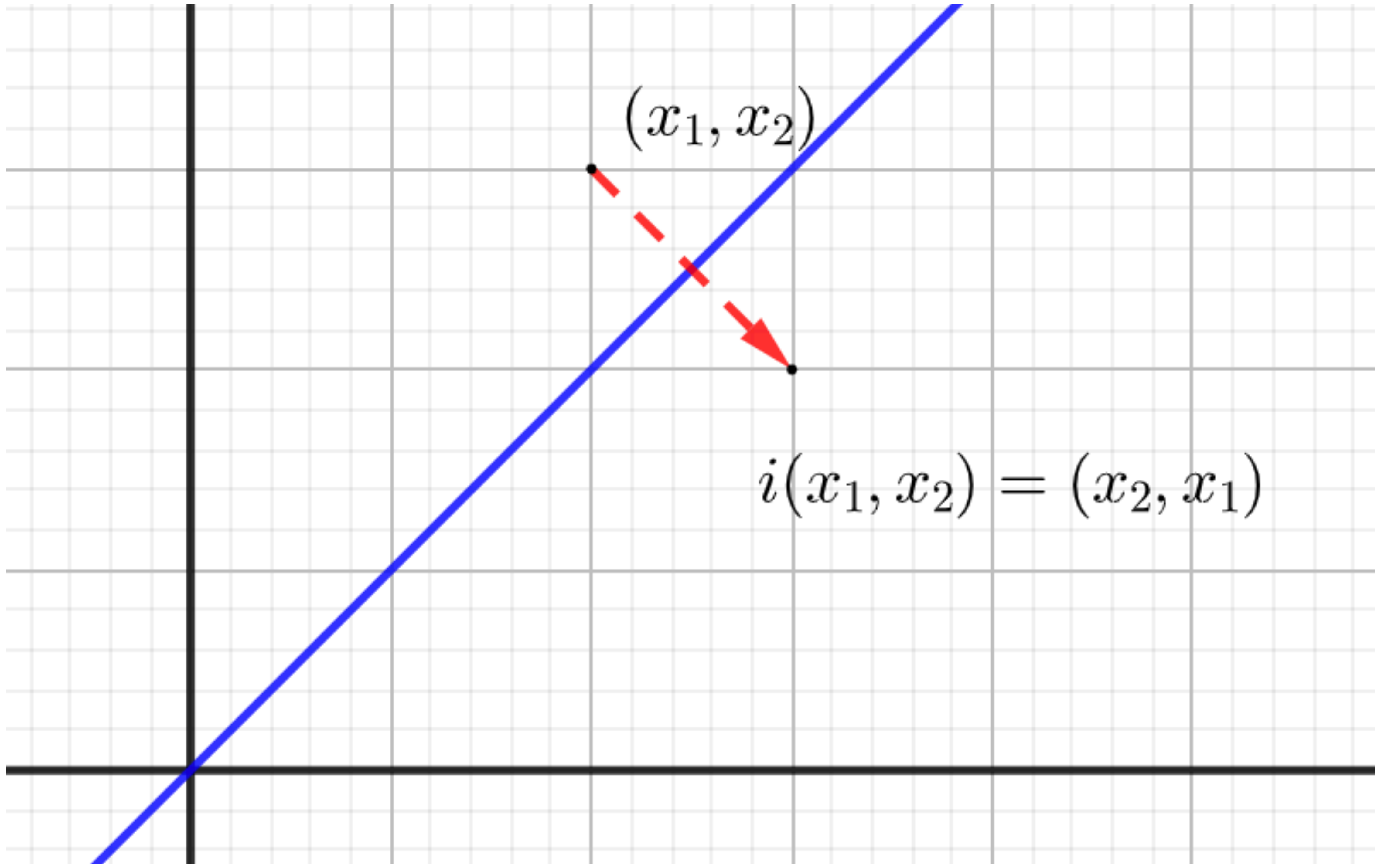
The function has the form

$h(x_1, x_2) = (-x_1, -x_2)$.

The matrix C has the form

$$C = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Reflection: $x_2 = x_1$



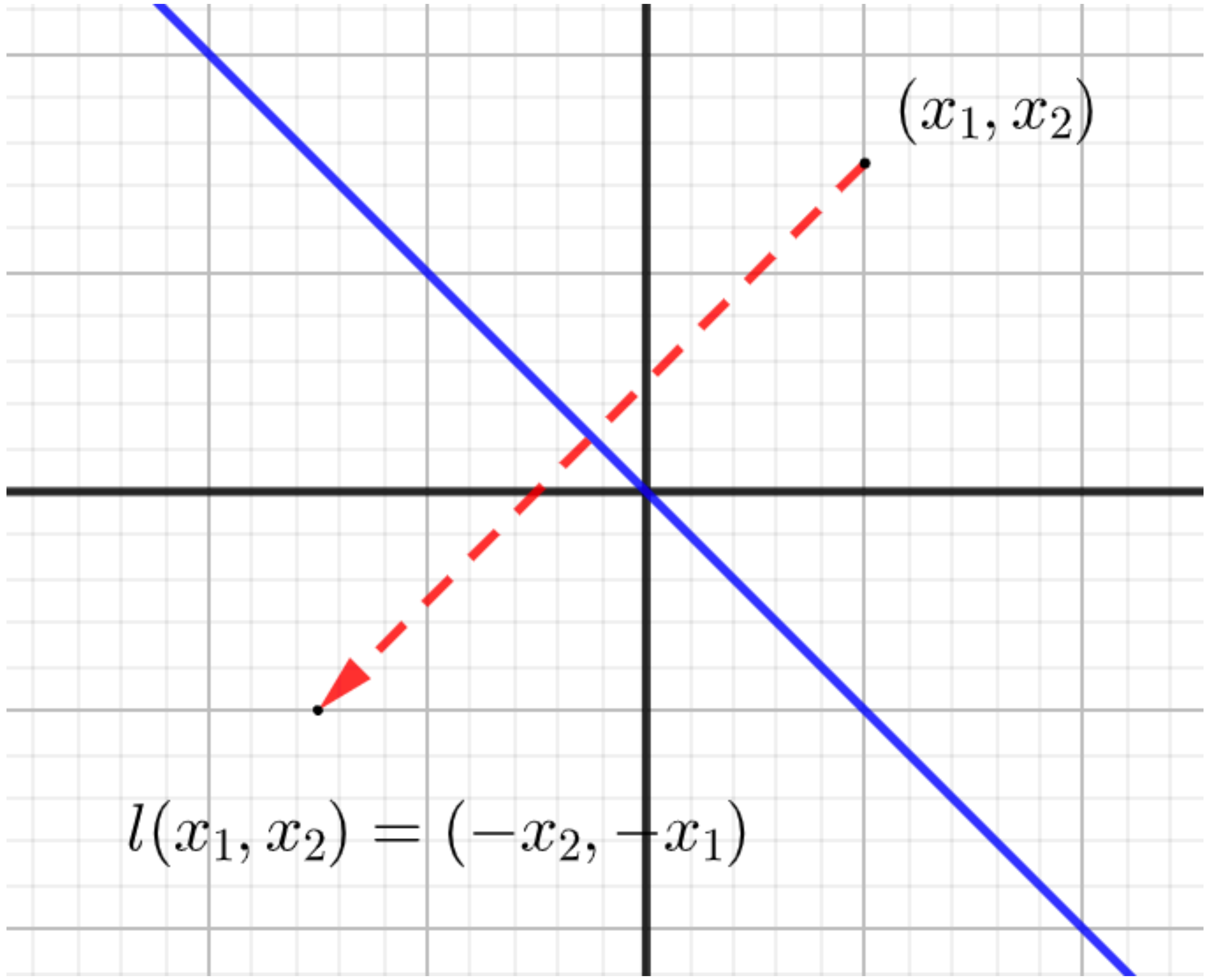
The function has the form

$i(x_1, x_2) = (x_2, x_1)$.

The matrix D has the form

$$D = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Reflection: $x_2 = -x_1$

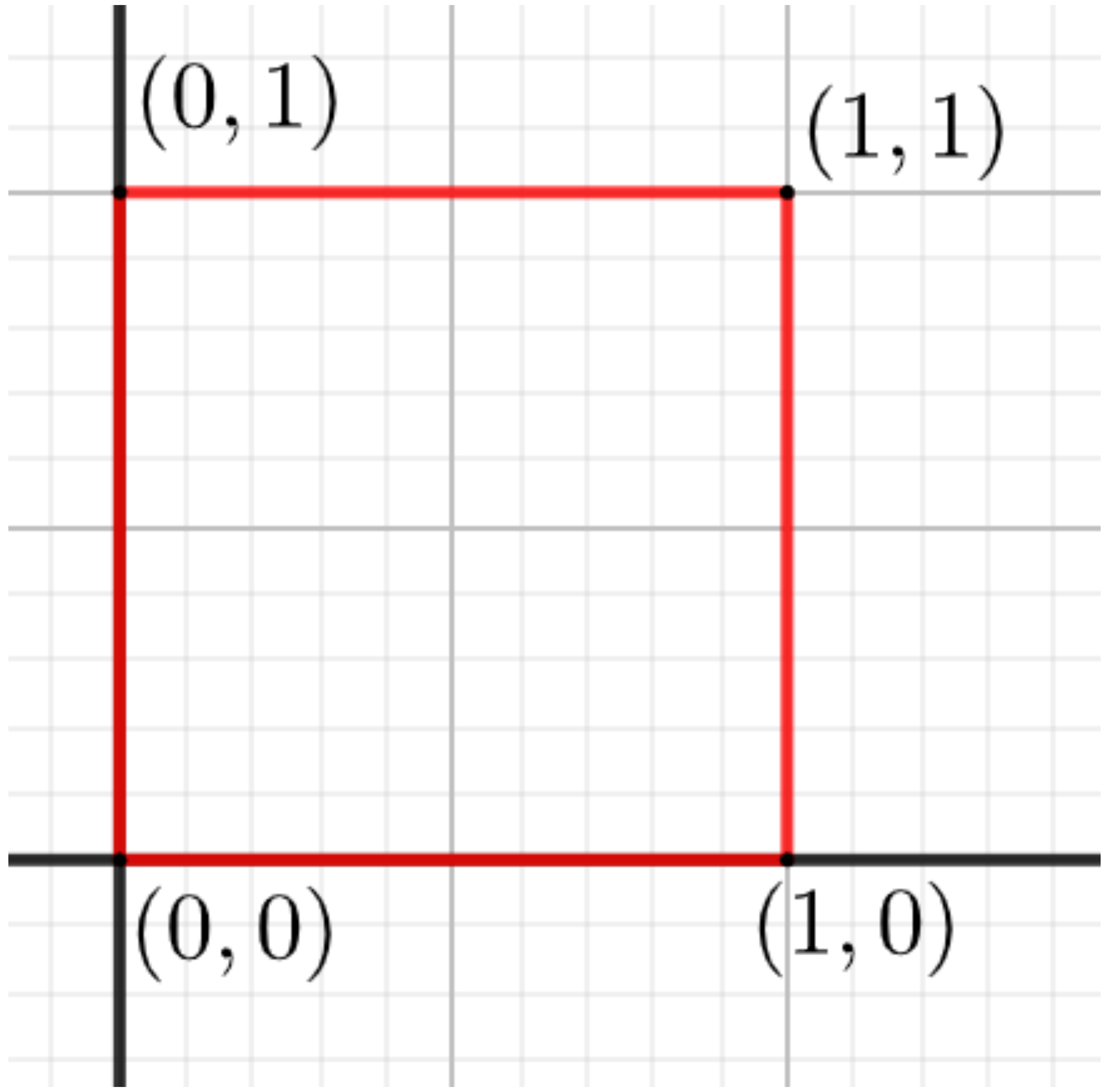


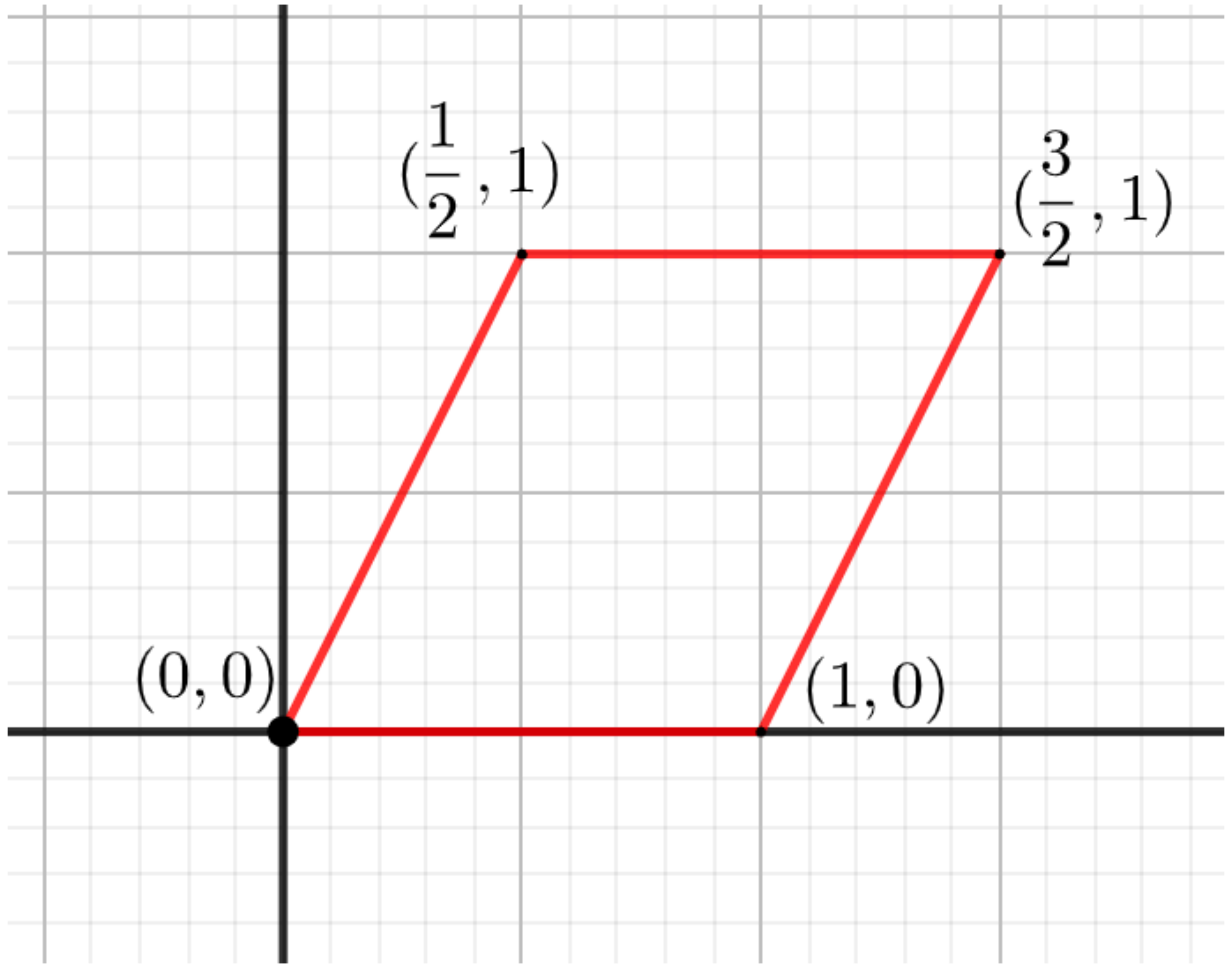
The function has the form
$l(x_1, x_2) = (-x_2, -x_1)$.

The matrix E has the form

$$E = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Shear transformations: horizontal shear

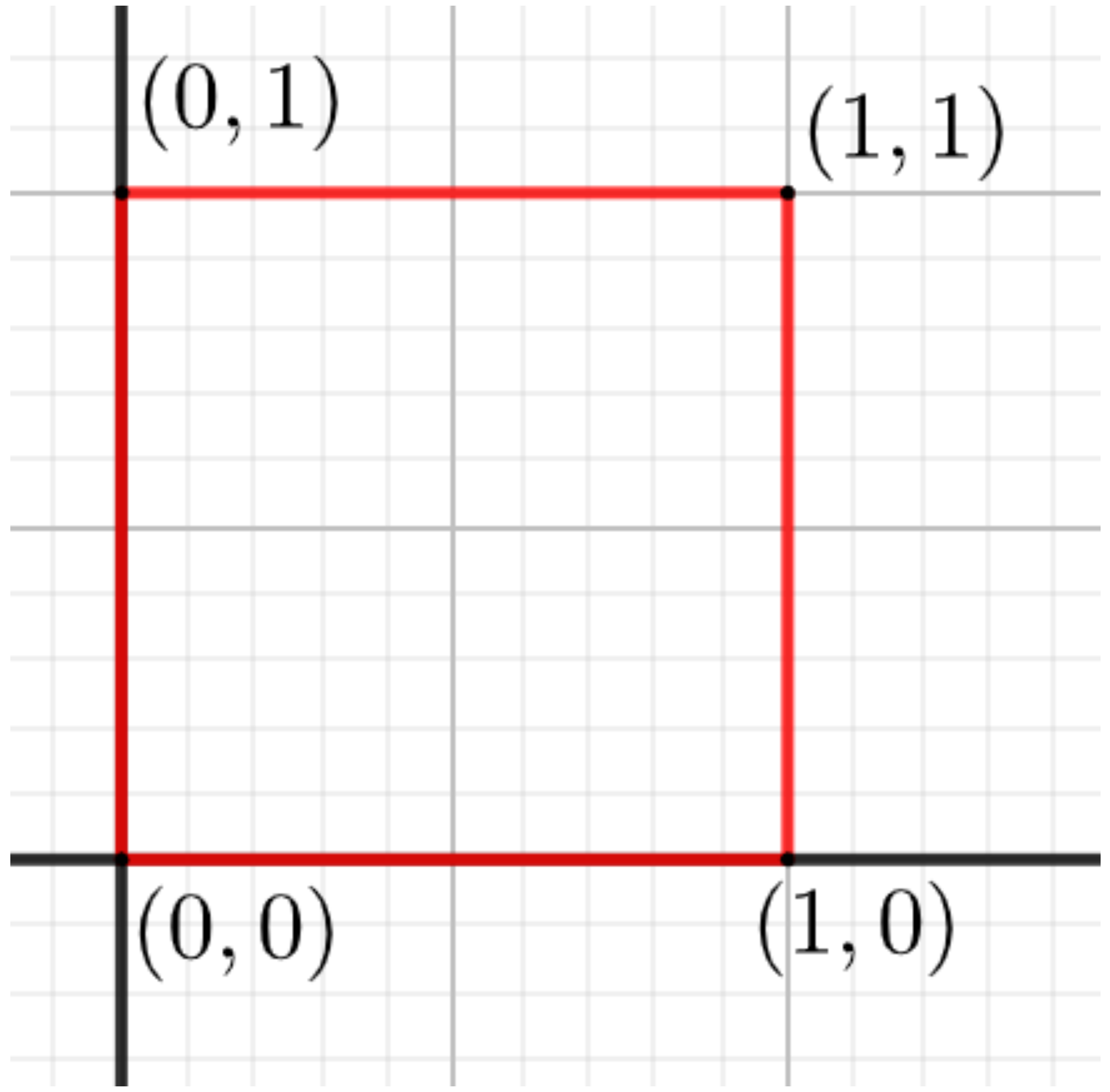


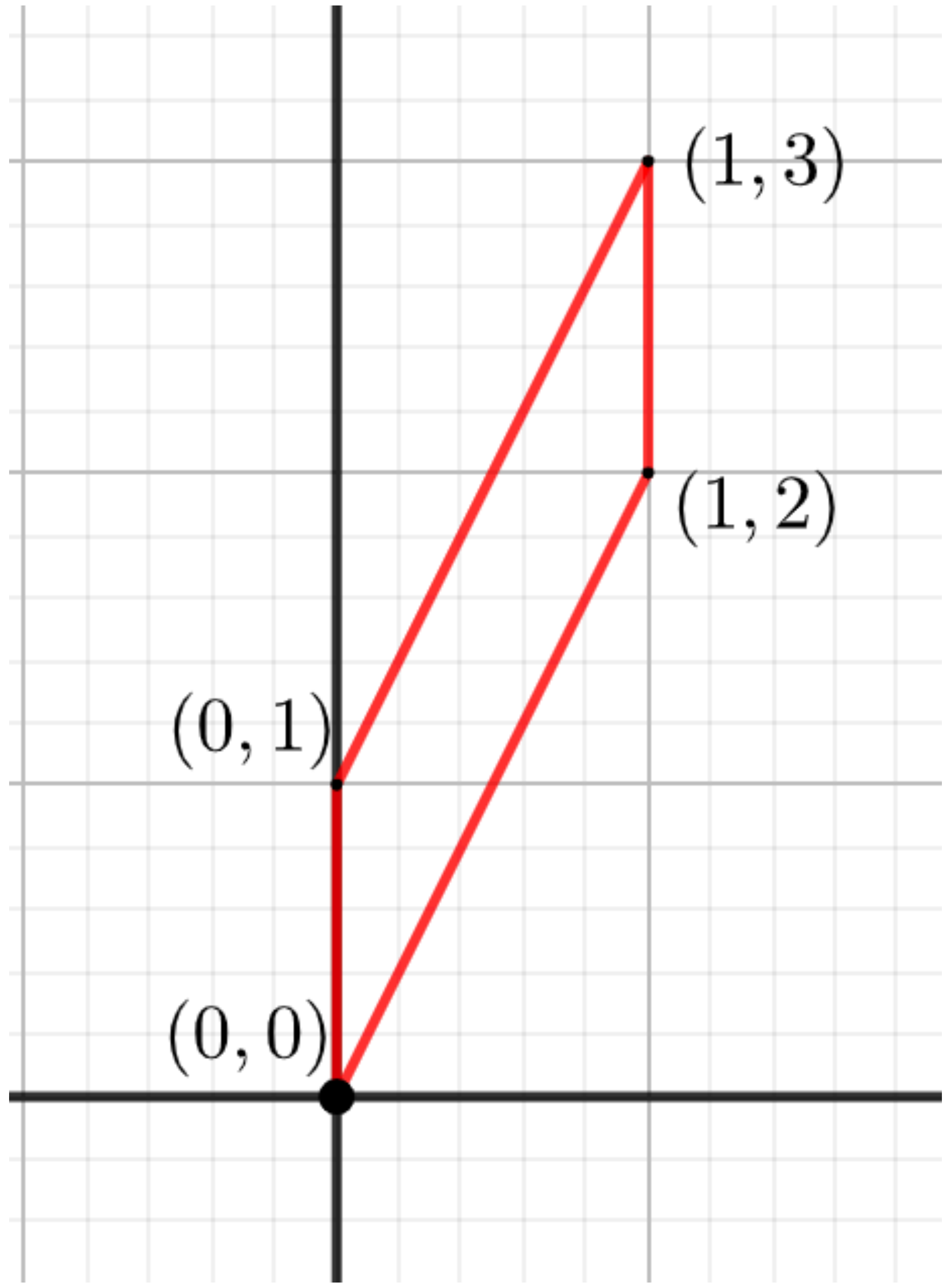


A horizontal shear will map horizontal lines to horizontal lines and vertical lines to oblique lines.

The generic transformation has the form $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$. In the picture the matrix is $A = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$

Shear transformations: vertical shear





A vertical shear maps horizontal lines to oblique lines and vertical lines to to vertical lines. . Generic form is $A = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$, while in the picture it is $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$.

Affine transformations

An **affine transformation** on $\mathbb{R}^m$ is a function $F : \mathbb{R}^m \to \mathbb{R}^n$ of the form $F(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ for $A \in \mathbb{R}^{n \times m}$ and $\mathbf{b} \in \mathbb{R}^n$. It is therefore a function that differs from a linear function by an additive term $\mathbf{b}$.

Affine transformations: translation

The geometric interpretation of the additive term is a translation. In the picture, we can see the action of the function

$$F(\mathbf{x}) = \mathbf{x} + \mathbf{b} \text{ for } \mathbf{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$