Lecture 10: Quadratic forms and Singular Value Decomposition

Francesco Preta

08/06/2020

Quadratic forms

Definition of quadratic form

A **quadratic form** in $\mathbb{R}^n$ is an expression of the type

$$Q(\mathbf{x}) = \langle \mathbf{x}, A\mathbf{x} \rangle$$

for a symmetric matrix $A \in \mathbb{R}^{n \times n}$.

Example:

$$A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

Example: diagonal quadratic form

A quadratic form has a particularly good expression if A is a diagonal matrix.

$$A = \begin{bmatrix} 9 & 0 \\ 0 & -4 \end{bmatrix}$$

Classification of quadratic forms

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form, then Q is said to be

- **positive definite** if $Q(\mathbf{x}) > 0$ for every $\mathbf{x} \neq \mathbf{0}$.
- **semi-positive definite** if $Q(\mathbf{x}) \geq 0$ for every $\mathbf{x} \in \mathbb{R}^n$.
- **negative definite** if $Q(\mathbf{x}) < 0$ for every $\mathbf{x} \neq \mathbf{0}$.
- **semi-negative definite** if $Q(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in \mathbb{R}^n$.
- **indefinite** otherwise.

Examples:

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}, \; A_2 = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

How to classify non-diagonal quadratic forms

Generally, if Q is a quadratic form associated to a non-diagonal matrix A, it's harder to say whether Q is positive (semi)-definite, negative (semi)-definite or indefinite just by looking at the explicit expression. However, since we are dealing with symmetric matrices, we can use its spectral property to change variables and diagonalize the quadratic form.

Example:

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

Change of variable

If U is the orthogonal matrix having unit eigenvectors of A as columns, then we can consider the change of variable $\mathbf{y} = U^T\mathbf{x}$, to have

$$Q(\mathbf{x}) = \langle \mathbf{x}, A\mathbf{x} \rangle = \langle \mathbf{x}, U\Lambda U^T\mathbf{x} \rangle = \langle U^T\mathbf{x}, \Lambda U^T\mathbf{x} \rangle = \langle \mathbf{y}, \Lambda\mathbf{y} \rangle$$

this last inner product can be expressed as $\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$, for which the classification is possible on the basis of the coefficients $\lambda_1, \ldots, \lambda_n$.

Classification theorem for quadratic forms

Theorem

Let A be a symmetric matrix associated to the quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$. Then

- Q is positive (resp. negative) definite if all its eigenvalues are strictly greater (resp. less) than 0.
- Q is semi-positive (resp. semi-negative) definite if all its eigenvalues are greater (resp. less) than or equal to 0.
- Q is indefinite if it has eigenvalues with different signs.

Example:

$$A = \begin{bmatrix} 9 & -4 & 4 \\ -4 & 7 & 0 \\ 4 & 0 & 11 \end{bmatrix}$$

Geometric interpretation of quadratic forms in $\mathbb{R}^2$

Consider a form in $Q : \mathbb{R}^2 \to \mathbb{R}$ having matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Then the form will be

$$Q(\mathbf{x}) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

this can be seen as the graph of a function in $\mathbb{R}^2$, thus corresponding to a quadric in $\mathbb{R}^3$, having a conic as a level set.

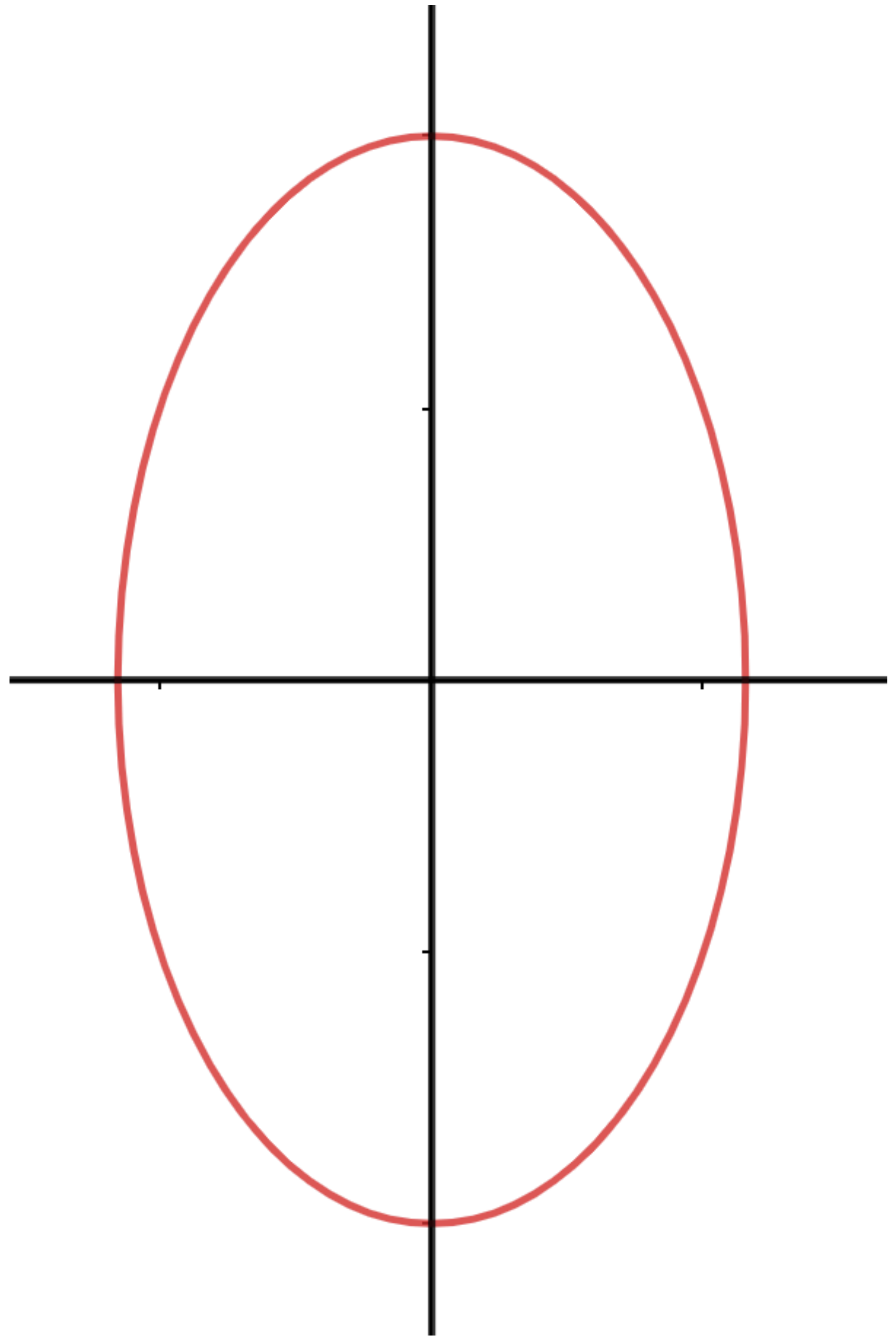
Classifying quadrics and conics

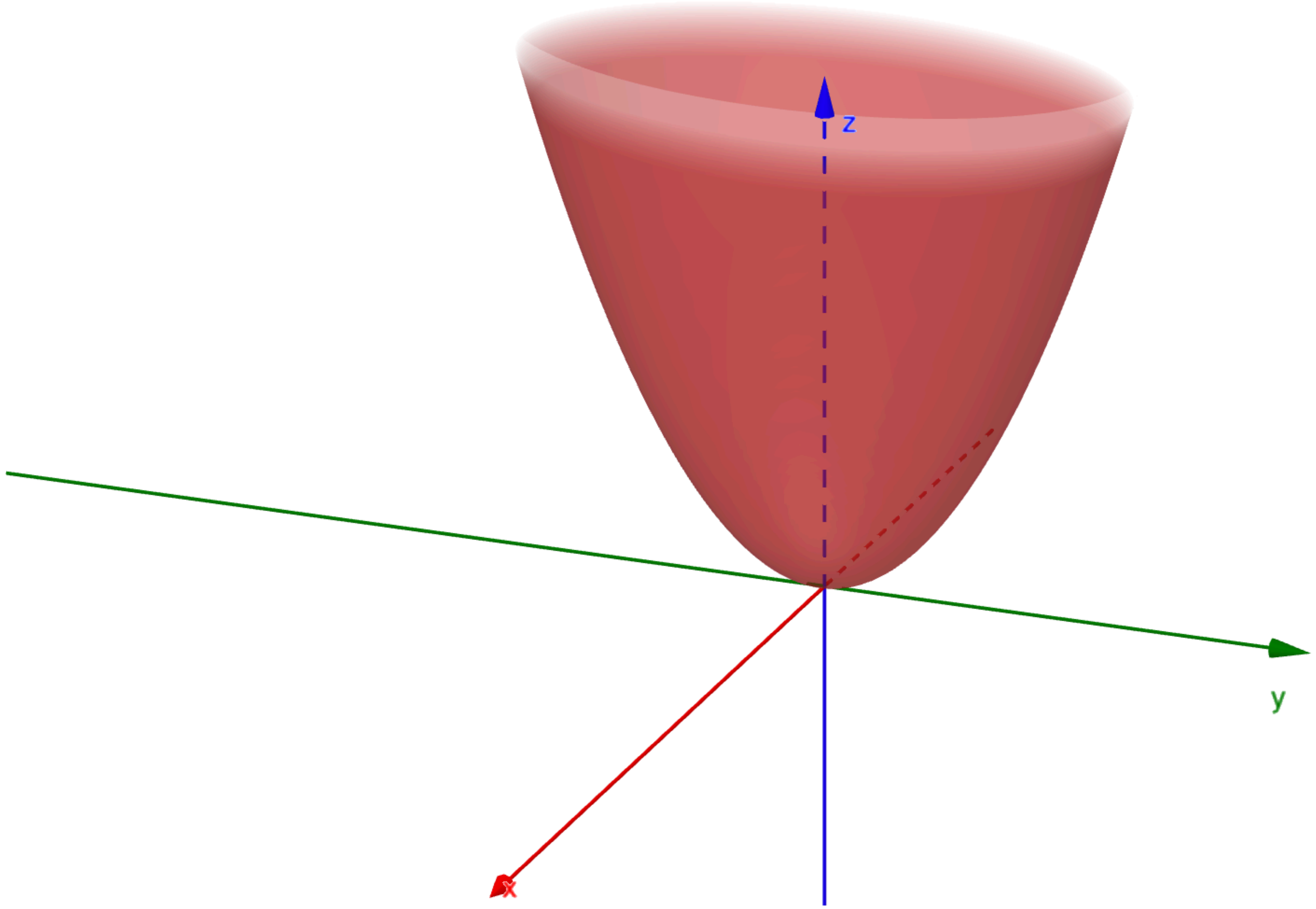
Through a change of variables we can diagonalize the quadratic form to obtain a form$Q(\mathbf{y}) = \lambda_1 y_1^2 + \lambda_2 y_2^2$ which gives the following possibilities:

- If Q is positive or negative definite, its graph is a paraboloid in $\mathbb{R}^3$ and its level sets are ellipses, for the levels at which they're well-defined.
- If Q is semi-definite (but not positive or negative definite), then its graph is a parabolic cylinder and its level sets are parallel lines.
- If Q is indefinite, then its graph is a parabolic hyperboloid and its level sets are hyperbolas.

Example: paraboloid

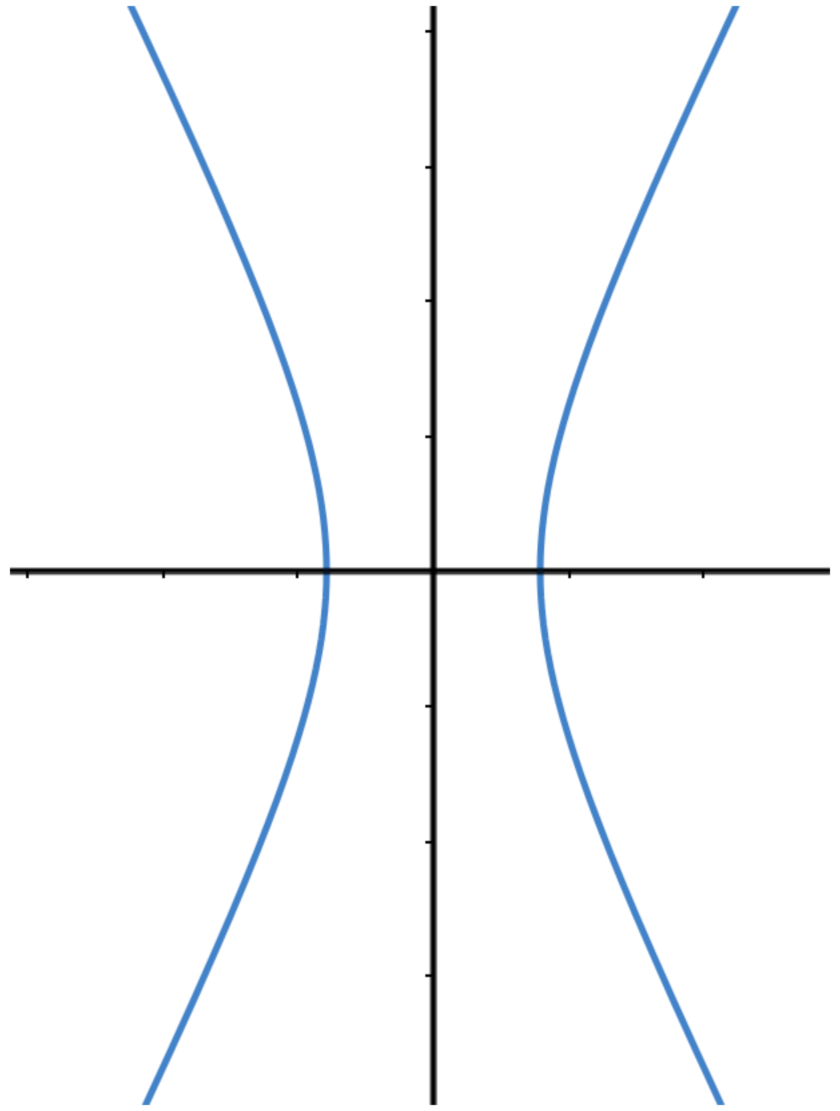
$z = 3x^2 + y^2$

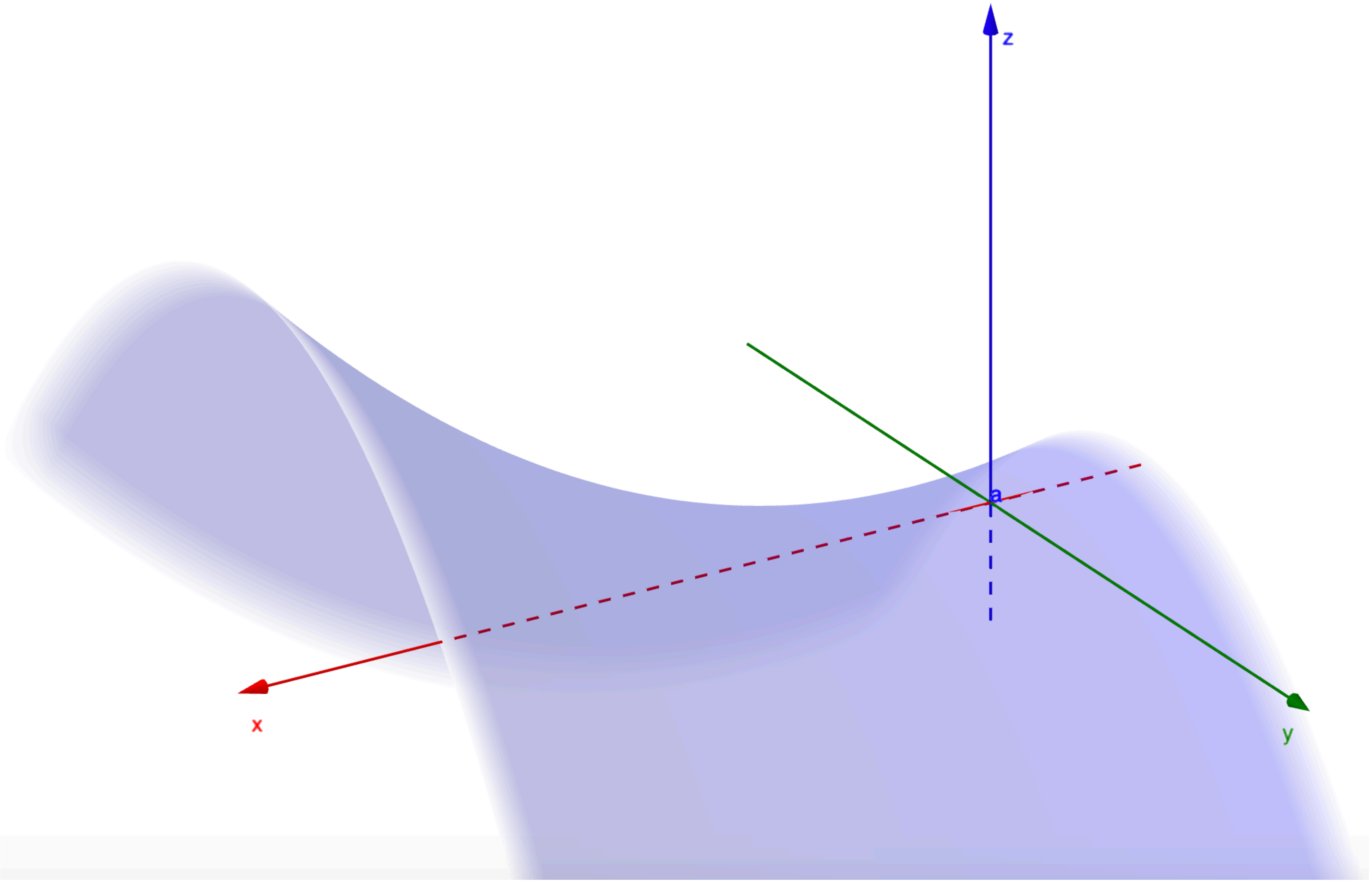




Example: hyperbolic paraboloid

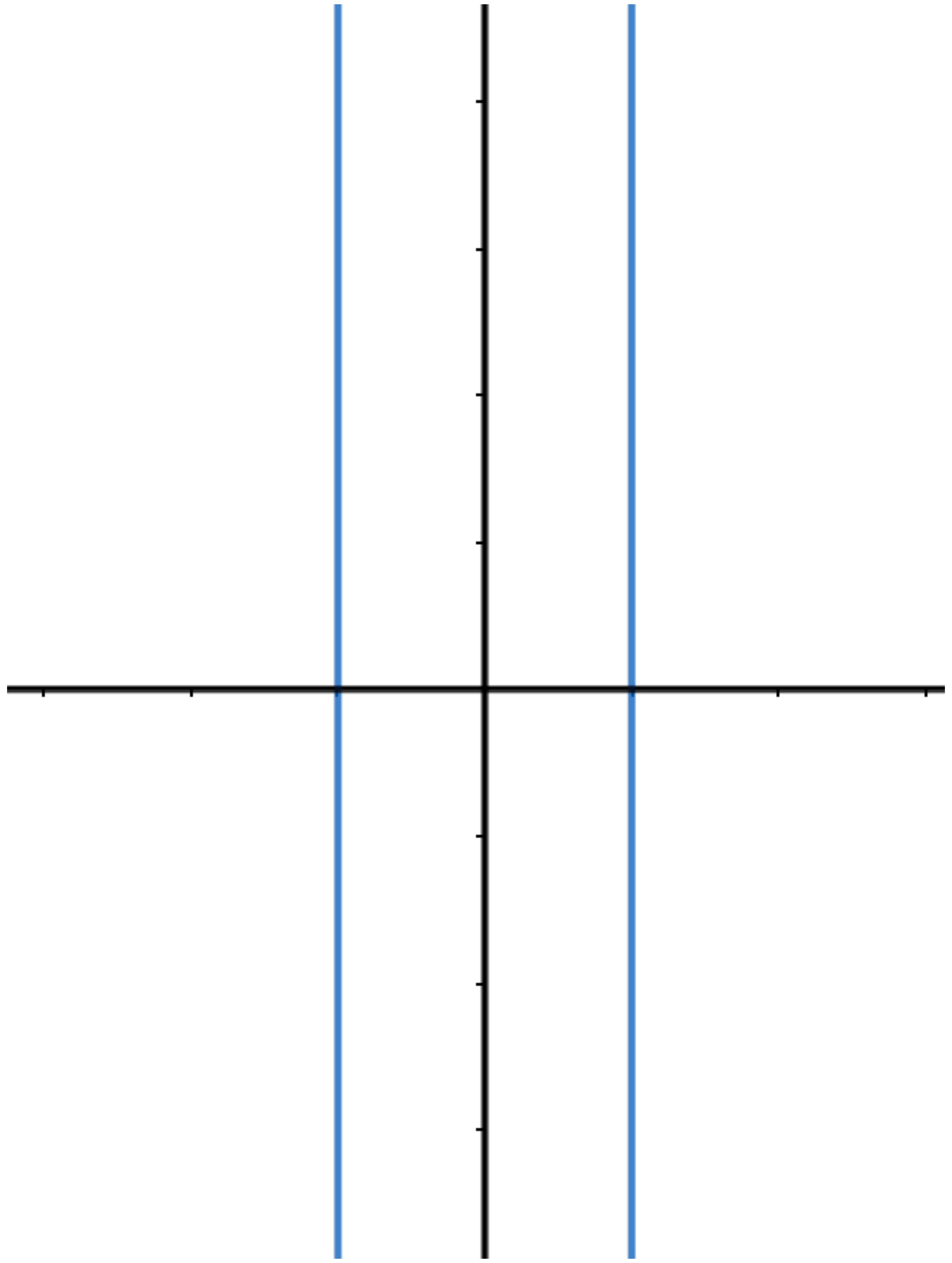
$$z = 0.4x^2 - 0.1y^2$$

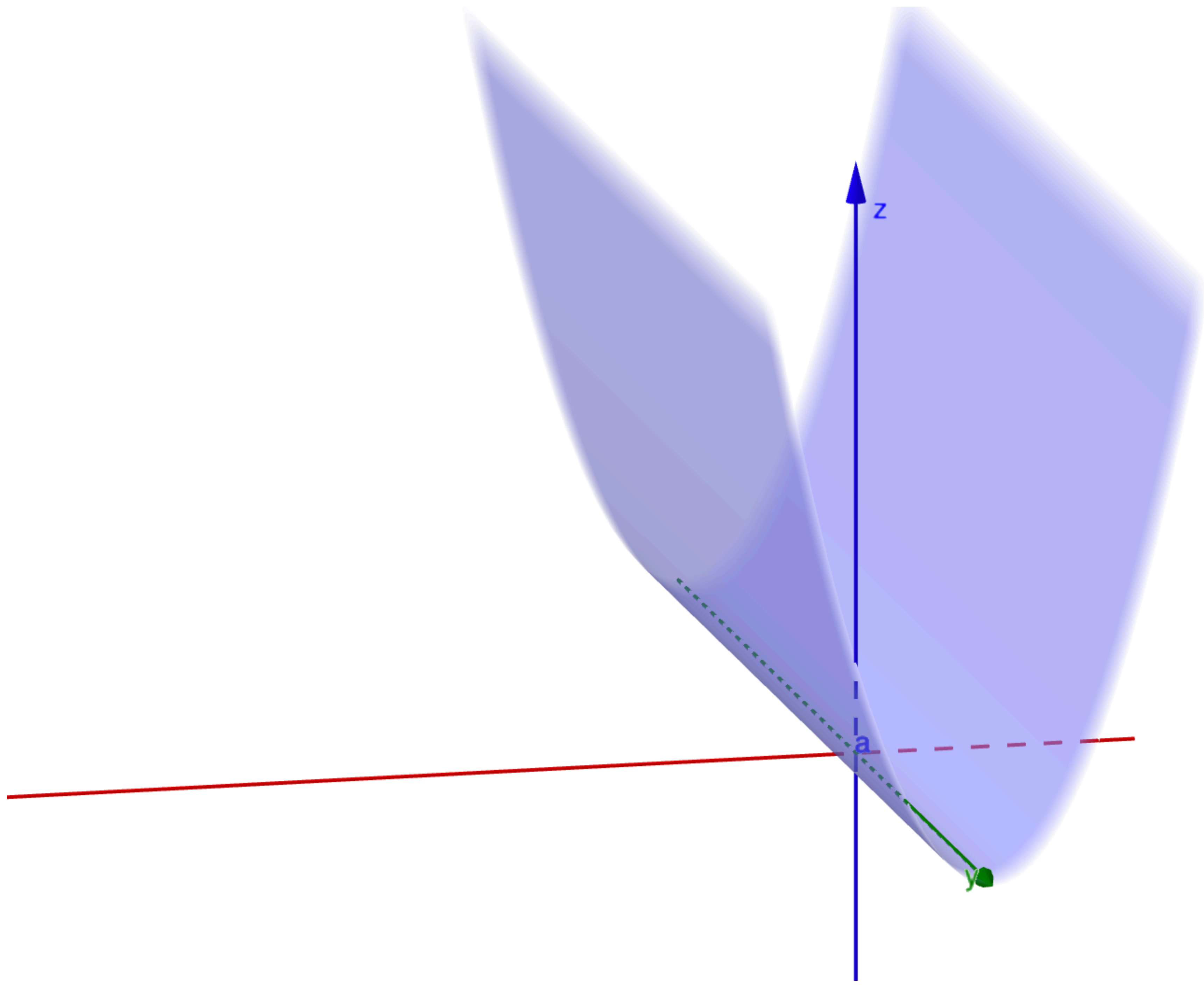




Example: parabolic cylinder

$z = x^2$





Classifying conics

The diagonalization procedure can also be used to classify a given conic whose form is not clear from the expression.

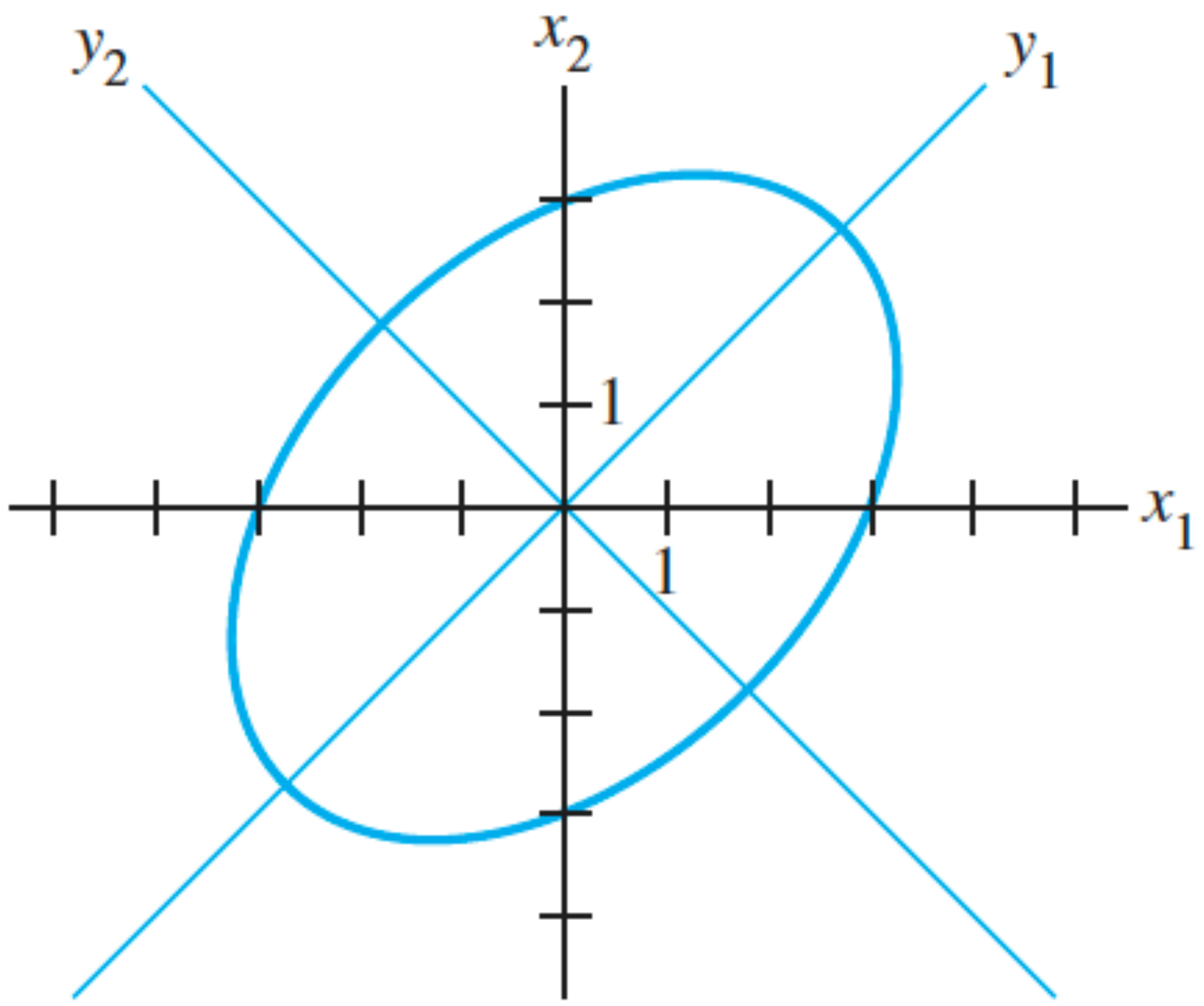
Example:

$$5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$$

Principal axis

In general, for a given conic $ax_1^2 + 2bx_1x_2 + cx_2^2 = d$, it is always possible to classify the conic by analyzing the eigenvalues of $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ and to find the **principal axis** corresponding to the directions of the eigenvectors. The transformation $\mathbf{y} = U^T\mathbf{x}$ corresponds to a transformation of $\mathbb{R}^2$ sending the principal axis to the coordinate axis.

Example:



(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$

Maximum and minimum of quadratic forms

If we look for optima for a quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ in the whole space $\mathbb{R}^n$, then we can find arbitrarily large values. However, we can restrict our optimization on the unit sphere, that is, on vectors of unit norm. In other words, we want to find

$M = \max_{||\mathbf{x}||=1} Q(\mathbf{x})$ and $m = \min_{||\mathbf{x}||=1} Q(\mathbf{x})$

Example:

$$Q(\mathbf{x}) = 9x_1^2 + 3x_2^2 - 2x_3^2$$

Eigenvalues as maximum and minimum

The same property can be applied to any quadratic form, up to diagonalization of its corresponding symmetric matrix.

Theorem

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form with corresponding symmetric matrix A and let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of A in decreasing order. Then

$$\lambda_1 = \max_{||\mathbf{x}||=1} \langle \mathbf{x}, A\mathbf{x} \rangle$$

$$\lambda_n = \min_{||\mathbf{x}||=1} \langle \mathbf{x}, A\mathbf{x} \rangle$$

Moreover, the point of maximum is the unitary eigenvector $\mathbf{v}_1$, while the point of minimum is the unitary eigenvector $\mathbf{v}_n$.

Proof:

Rayleigh quotient and generalization

Rayleigh quotient Theorem

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form with corresponding symmetric matrix A and let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of A in decreasing order. Then

$$\lambda_i = \max\{\langle \mathbf{x}, A\mathbf{x}\rangle \mid ||\mathbf{x}|| = 1, \langle \mathbf{x}, \mathbf{v}_j\rangle = 0 \text{ for all } j < i\}$$

Singular Value Decomposition

Singular Value Decomposition

The singular value decomposition is one of the most important decompositions in applied linear algebra and it's a generalization of diagonalization for square matrices.

It generalizes the idea of eigenvalues as optima, but for the maximization of the norm $||A\mathbf{x}||$, even if A is not a square matrix.

Example:

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

the linear operator $A : \mathbb{R}^3 \to \mathbb{R}^2$ send spheres in $\mathbb{R}^3$ to ellipsis in $\mathbb{R}^2$. We are interested in understanding the length of the principal axis, corresponding to the maximum value of $||A\mathbf{x}||$ for $\mathbf{x} \in \mathbb{R}^3$ of unit norm and to the length of its orthogonal axis.

Maximum norm: eigenvalues of $A^T A$

Since

$$||A\mathbf{x}||^2 = \langle A\mathbf{x}, A\mathbf{x} \rangle = \langle \mathbf{x}, A^T A\mathbf{x} \rangle$$

the maximum value of $||A\mathbf{x}||$, will therefore be given by the square root of the maximum eigenvalue of $A^T A$.

Example:

In our case $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Definition: Singular Values

Let A be a $m \times n$ matrix. The **singular values** of A are the square roots of the eigenvalues of the $n \times n$ matrix A^TA, that is, if $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of A^TA, then the singular values are $\sigma_i = \sqrt{\lambda_i}$ for $i = 1,\ldots,n$.

Singular values are well-defined

The square root is well-defined in $\mathbb{R}$, since $A^T A$ is positive semi-definite, therefore all its eigenvalues are greater or equal to 0. This happens because

$$\langle \mathbf{x}, A^T A \mathbf{x} \rangle = \langle A\mathbf{x}, A\mathbf{x} \rangle = ||A\mathbf{x}||^2 \geq 0$$

Point of maximum

If $\mathbf{v}_i$ is the unitary eigenvector of $A^T A$ corresponding to the eigenvalue λ_i, then

$$\sigma_i = \sqrt{\lambda_i} = \sqrt{\langle \mathbf{v}_i, A^T A \mathbf{v}_i \rangle} = \sqrt{\langle A\mathbf{v}_i, A\mathbf{v}_i \rangle} = ||A\mathbf{v}_i||$$

Orthogonality is preserved by A

Theorem

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be an orthonormal basis of eigenvectors for $A^T A$ in $\mathbb{R}^n$ and let $r \leq n$ be such that A has r nonzero singular values. Then $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$ are an orthogonal basis for Im(A).

Proof:

Example:

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

The minor principal axis of the ellipsis is given by the image of the second eigenvector of $A^T A$.

Singular value decomposition

Given a matrix A we can always find a (left) singular value decomposition:

Theorem (Singular Value Decomposition)

Let A be a $m \times n$ matrix having $r \leq \min\{m, n\}$nonzero singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$. Then there exists an orthogonal $n \times n$ matrix V and an orthogonal $m \times m$ matrix U such that

$$A = U\Sigma V^T$$

where Σ is a $m \times n$ matrix having σ_i on the i-th element of its diagonal and 0 everywhere else.

Example:

With the previous example $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$, let's follow the construction.

Process to find SVD: Step 1

Step 1: find the eigenvalues and the unitary eigenvectors of the symmetric matrix A^TA. These will be $\lambda_1 \geq \lambda_2 \geq \lambda_r > 0$ and $\lambda_{r+1} = \ldots = \lambda_n = 0$ and the corresponding $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Process to find SVD: Step 2

Step 2: construct Σ and V. Σ is the matrix such that $\Sigma_{i,i} = \sigma_i = \sqrt{\lambda_i}$ for $i = 1,\ldots,r$ and 0 otherwise, while V is the matrix having $\mathbf{v}_i$ as i-th column.

Process to find SVD: Step 3

Step 3: for $i = 1,\ldots,r$, find $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.

Process to find SVD: Step 4

Step 4: complete $\mathbf{u}_1, \ldots, \mathbf{u}_r$ to a basis or $\mathbb{R}^m$ and write U the matrix having $\mathbf{u}_i$ as i-th column.

Why does it work?

The reason why this happens is because we have

$$U\Sigma = \begin{bmatrix} A\mathbf{v}_1 & \dots & A\mathbf{v}_r & \mathbf{0} \end{bmatrix} = AV$$

$$U\Sigma V^T = AVV^T = A$$

Example:

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

Exercises:

Classify the following quadratic forms:

- $Q(\mathbf{x}) = 2x_1^2 - 4x_1x_2 - x_2^2$
- $Q(\mathbf{x}) = -x_1^2 - 2x_1x_2 - x_2^2$
- $Q(\mathbf{x}) = x_1^2 - 2x_1x_2 + x_2^2$

For each of these forms, what is the conic representing $Q(\mathbf{x}) = 1$ in $\mathbb{R}^2$?

Exercises:

Find the SVD of

$$A = \begin{bmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$