

Symmetric Matrices

Geometric Algorithms

Lecture 25

CAS CS 132

Recap Problem

$$\{(0,3), (1,1), (-1,1), (2,3)\}$$

Find the matrices X as in the previous example to find the least squares best fit parabola and the least squares best fit cubic for this dataset.

$$\{(0,3), (1,1), (-1,1), (2,3)\}$$

Answer

Objectives

1. Talk about about symmetric matrices and eigenvalues.
2. Describe an application to constrained optimization problems.

Keywords

linear models

design matrices

general linear regression

symmetric matrices

the spectral theorem

orthogonal diagonalizability

quadratic forms

definiteness

constrained optimization

Symmetric Matrices

Recall: Symmetric Matrices

Definition. A square matrix A is **symmetric** if $A^T = A$.

Orthogonal Eigenvectors

$$A\vec{v} = \lambda\vec{v} \text{ for some } \lambda \in \mathbb{R}$$
$$\vec{v} \neq 0$$

Theorem. For a symmetric matrix A , if \mathbf{u} and \mathbf{v} are eigenvectors for *distinct* eigenvalues, then \mathbf{u} and \mathbf{v} are orthogonal.

Verify:

Recall: Diagonalizable Matrices

Recall: Diagonalizable Matrices

Definition. A matrix A is **diagonalizable** if it is similar to a diagonal matrix.

Recall: Diagonalizable Matrices

Definition. A matrix A is **diagonalizable** if it is similar to a diagonal matrix.

There is an invertible matrix P and diagonal matrix D such that $A = PDP^{-1}$.

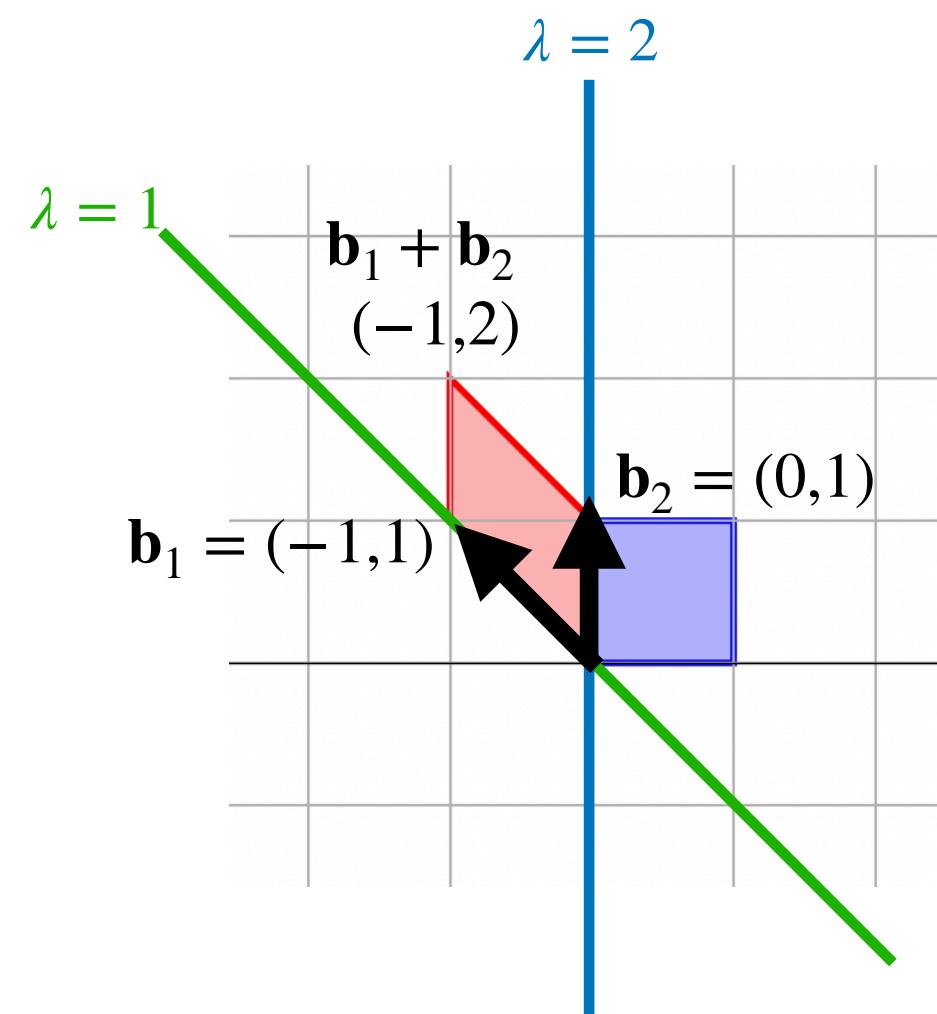
Recall: Diagonalizable Matrices

Definition. A matrix A is **diagonalizable** if it is similar to a diagonal matrix.

There is an invertible matrix P and diagonal matrix D such that $A = PDP^{-1}$.

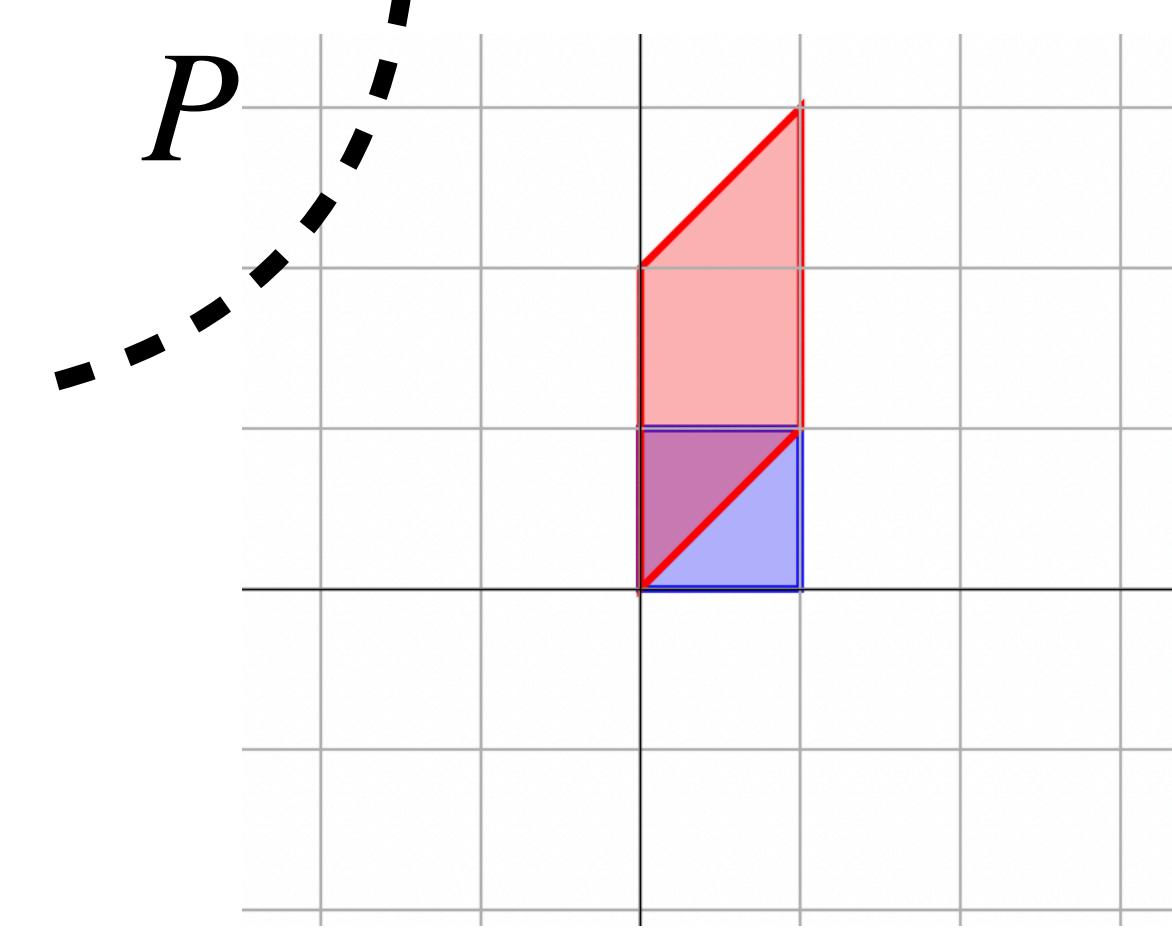
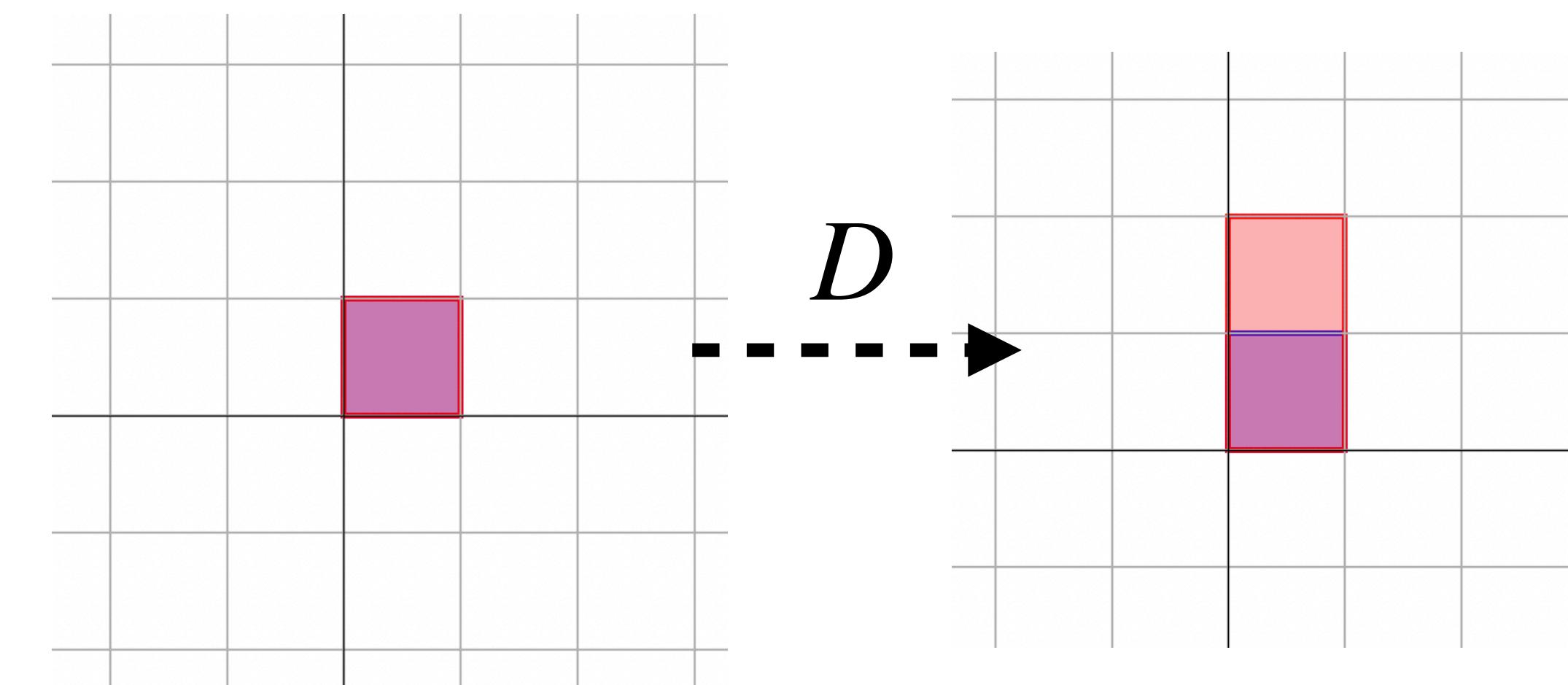
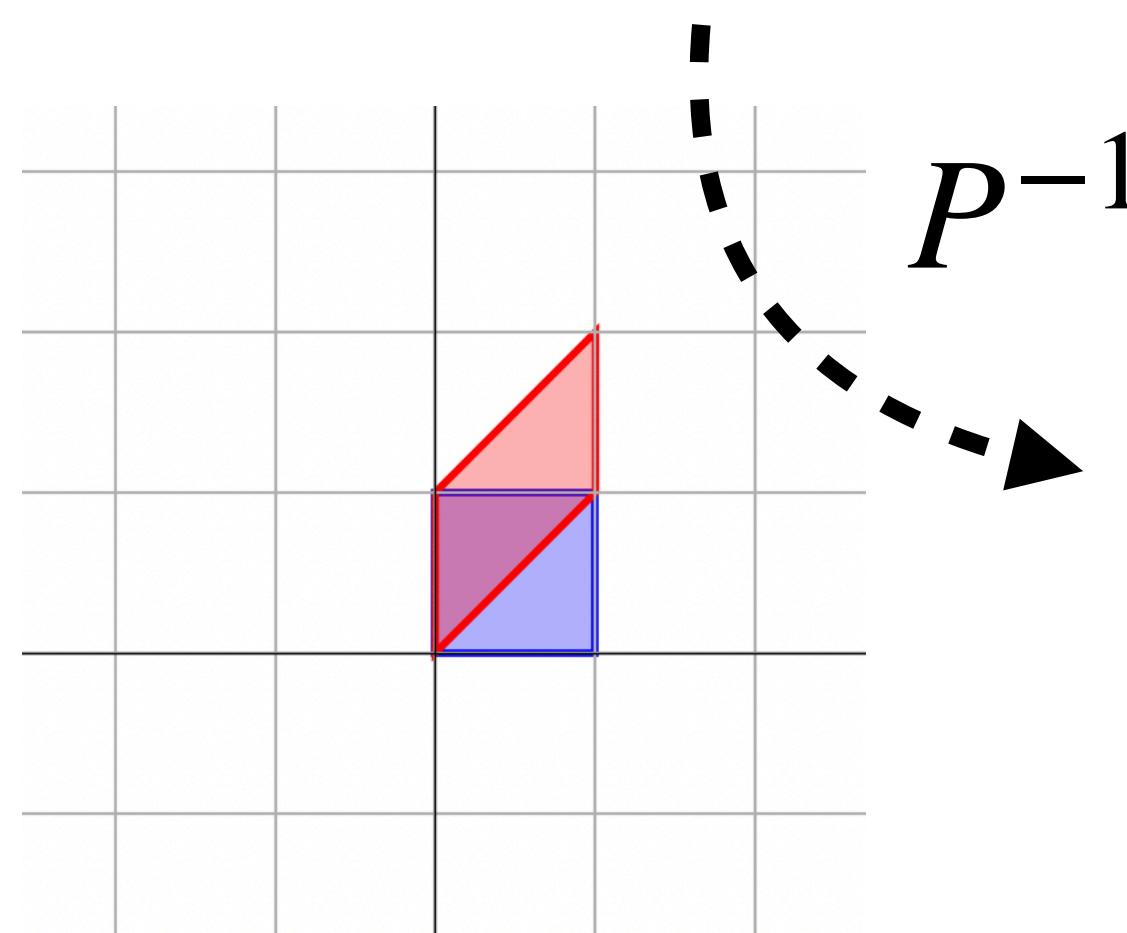
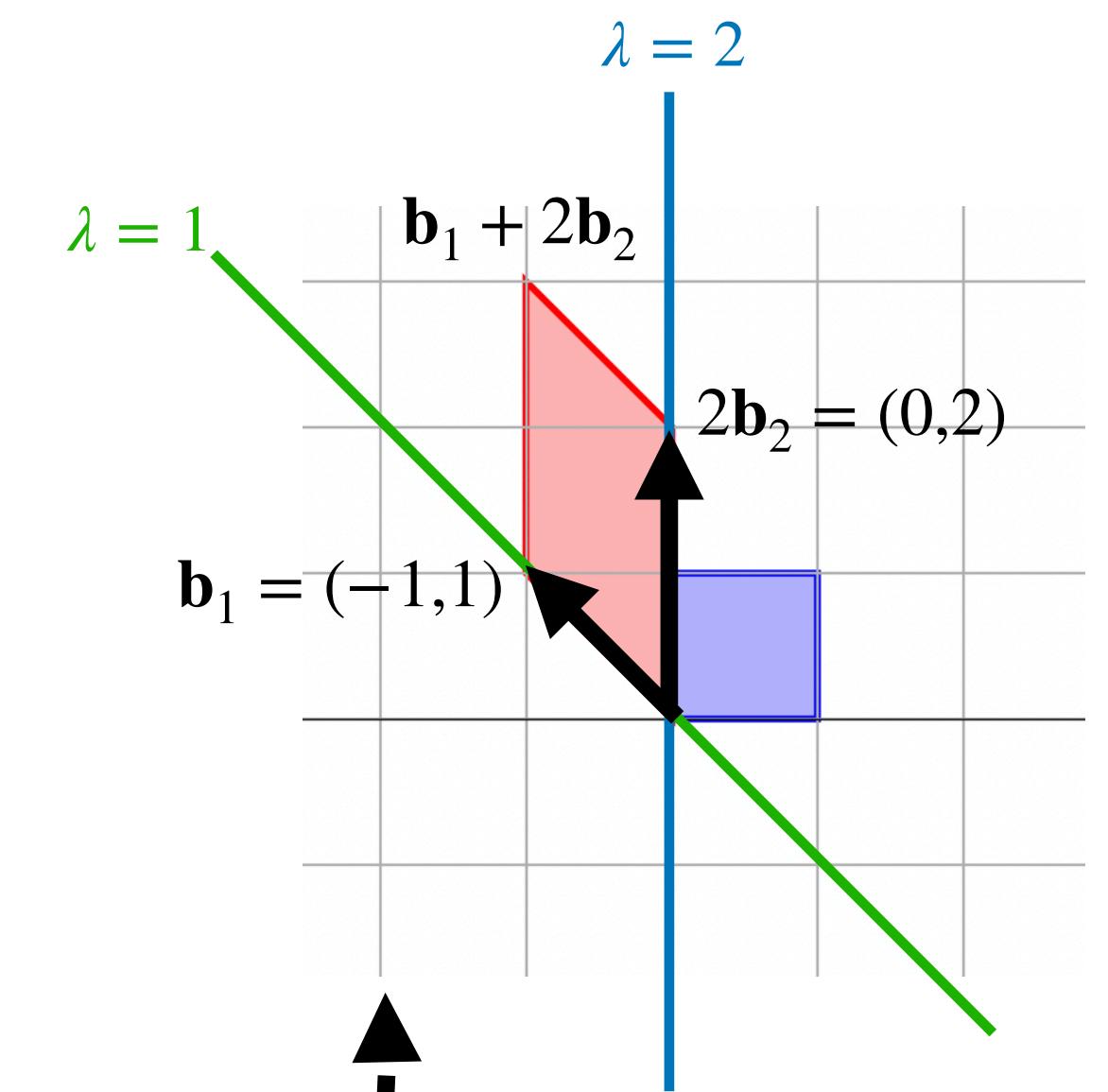
Diagonalizable matrices are the same as scaling matrices up to a change of basis.

Recall: The Picture



$$A = PDP^{-1}$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}$$



Recall: The Diagonalization Theorem

$$A = PDP^{-1}$$

Recall: The Diagonalization Theorem

$$A = PDP^{-1}$$

Theorem. A is diagonalizable if and only if it has an eigenbasis.

Recall: The Diagonalization Theorem

$$A = PDP^{-1}$$

Theorem. A is diagonalizable if and only if it has an eigenbasis.

The idea:

Recall: The Diagonalization Theorem

$$A = \overset{\text{eigenbasis}}{P} D P^{-1}$$

Theorem. A is diagonalizable if and only if it has an eigenbasis.

The idea:

The columns of P form an eigenbasis for A .

Recall: The Diagonalization Theorem

$$A = \begin{matrix} \text{eigenbasis} \\ P \end{matrix} D \begin{matrix} P^{-1} \\ \text{eigenvalues} \end{matrix}$$

Theorem. A is diagonalizable if and only if it has an eigenbasis.

The idea:

The columns of P form an eigenbasis for A .

The diagonal of D are the eigenvalues for each column of P .

Recall: The Diagonalization Theorem

$$A = \begin{matrix} \text{eigenbasis} \\ P \end{matrix} D \begin{matrix} \\ \text{eigenvalues} \end{matrix} P^{-1}$$

Theorem. A is diagonalizable if and only if it has an eigenbasis.

The idea:

The columns of P form an eigenbasis for A .

The diagonal of D are the eigenvalues for each column of P .

The matrix P^{-1} is a change of basis to this eigenbasis of A .

The Spectral Theorem

Theorem. If A is symmetric, then it has an *orthonormal* eigenbasis.

(we won't prove this)

Symmetric matrices are diagonalizable.

But more than that, we can choose P to be *orthogonal*.

Recall: Orthonormal Matrices

Definition. A matrix is **orthonormal** if its columns form an orthonormal set.

The notes call a square orthonormal matrix an **orthogonal** matrix.

Recall: Inverses of Orthogonal Matrices

Theorem. If an $n \times n$ matrix U is orthogonal (square orthonormal) then it is invertible and

$$U^{-1} = U^T$$

Verify:

$$\left[\begin{array}{c} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{array} \right]^T = \left[\begin{array}{cccc} | & | & \cdots & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{array} \right] = \left[\begin{array}{c} \text{i}^{th} \text{ entry} \\ = \vec{u}_i \cdot \vec{u}_j \end{array} \right]$$

Orthogonal Diagonalizability

Definition. A matrix A is **orthogonally diagonalizable** if there is a diagonal matrix D and matrix P such that

$$A = PDP^T = PDP^{-1}$$

P must be an orthonormal matrix.

Symmetric matrices are
orthogonally diagonalizable

Orthogonal Diagonalizability and Symmetry

Fact. All orthogonally diagonalizable matrices are symmetric.

Verify:

Orthogonal Diagonalizability and Symmetry

Theorem. A matrix is orthogonally diagonalizable if and only if it is symmetric.

(We'll usually just use NumPy)

Practice Problem

Find an orthogonal diagonalization of $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

Answer

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Quadratic Forms

Quadratic Forms

Definition. A **quadratic form** is a function of variables x_1, \dots, x_n in which every term has degree two.

Examples:

Quadratic Forms and Symmetric Matrices

Fact. Every quadratic form can be represented as

$$\mathbf{x}^T A \mathbf{x}$$

where A is symmetric.

Example:

Example: Computing the Quadratic Form for a Matrix

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

This means, given a symmetric matrix A , we can compute its corresponding quadratic form:

Quadratic forms and Symmetric Matrices (Again)

Furthermore, we can generally say

$$\mathbf{x}^T A \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i x_j = \sum_{i=1}^n A_{ii} x_i^2 + \sum_{i \neq j} (A_{ij} + A_{ji}) x_i x_j$$

Verify:

A Slightly more Complicated Example

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ -1 & 0 & 5 \end{bmatrix}$$

Let's expand $\mathbf{x}^T A \mathbf{x}$:

Matrices from Quadratic Forms

$$Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$$

We can also go in the other direction. Let's express this as $\mathbf{x}^T A \mathbf{x}$:

How To: Matrices of Quadratic Forms

Problem. Given a quadratic form $Q(\mathbf{x})$, find the symmetric matrix A such that $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$.

Solution.

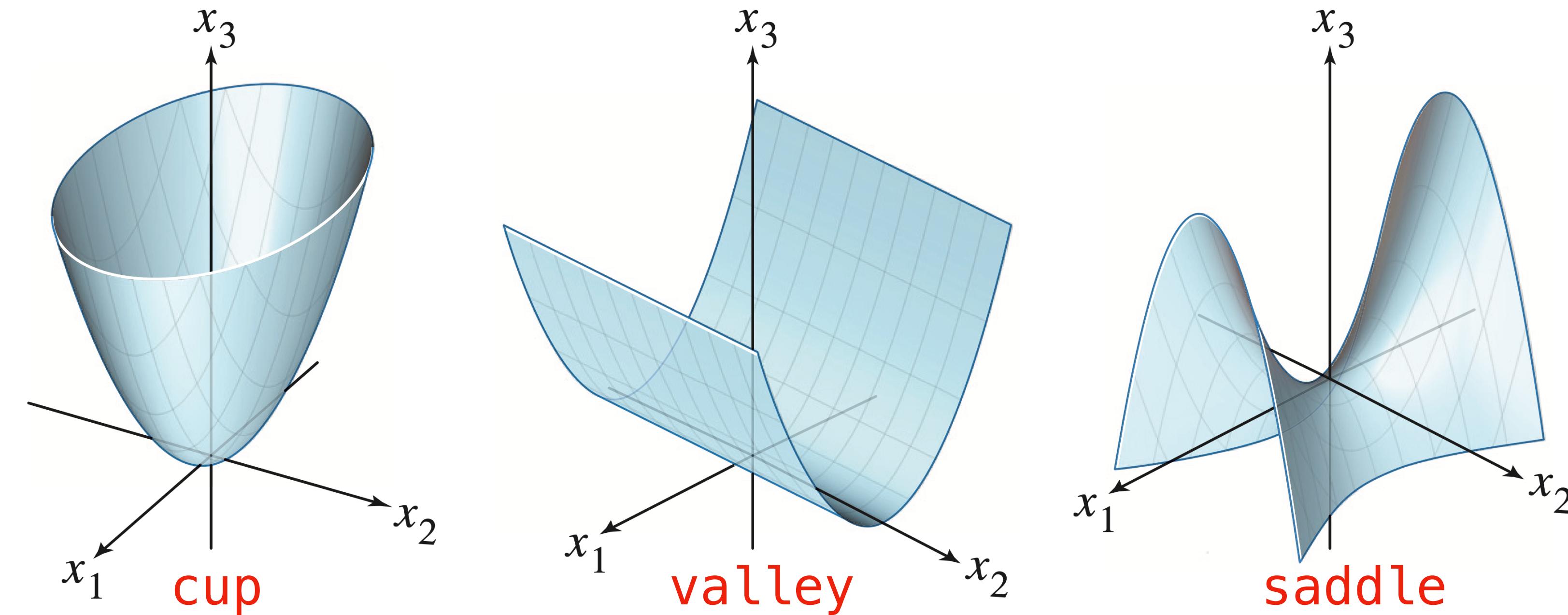
- » if $Q(\mathbf{x})$ has the term αx_i^2 then $A_{ii} = \alpha$
- » if $Q(\mathbf{x})$ has the term $\alpha x_i x_j$, then $A_{ij} = A_{ji} = \frac{\alpha}{2}$

Practice Problem

$$Q(x_1, x_2, x_3, x_4) = x_1^2 + 3x_2^2 - 2x_3x_4 - 6x_4^2 + 7x_1x_3$$

Find the symmetric matrix A such that $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$.

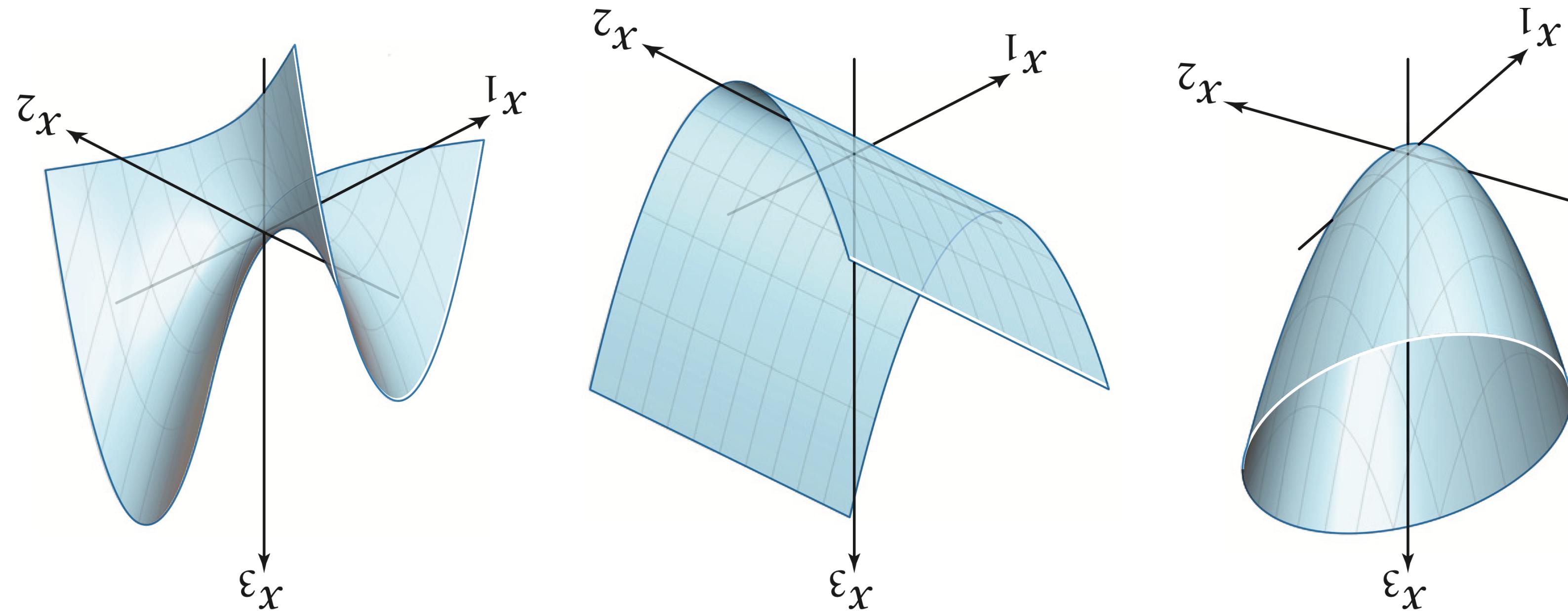
Shapes of Quadratic Forms



There are essentially three possible shapes (six if you include the negations).

Can we determine what shape it will be mathematically?

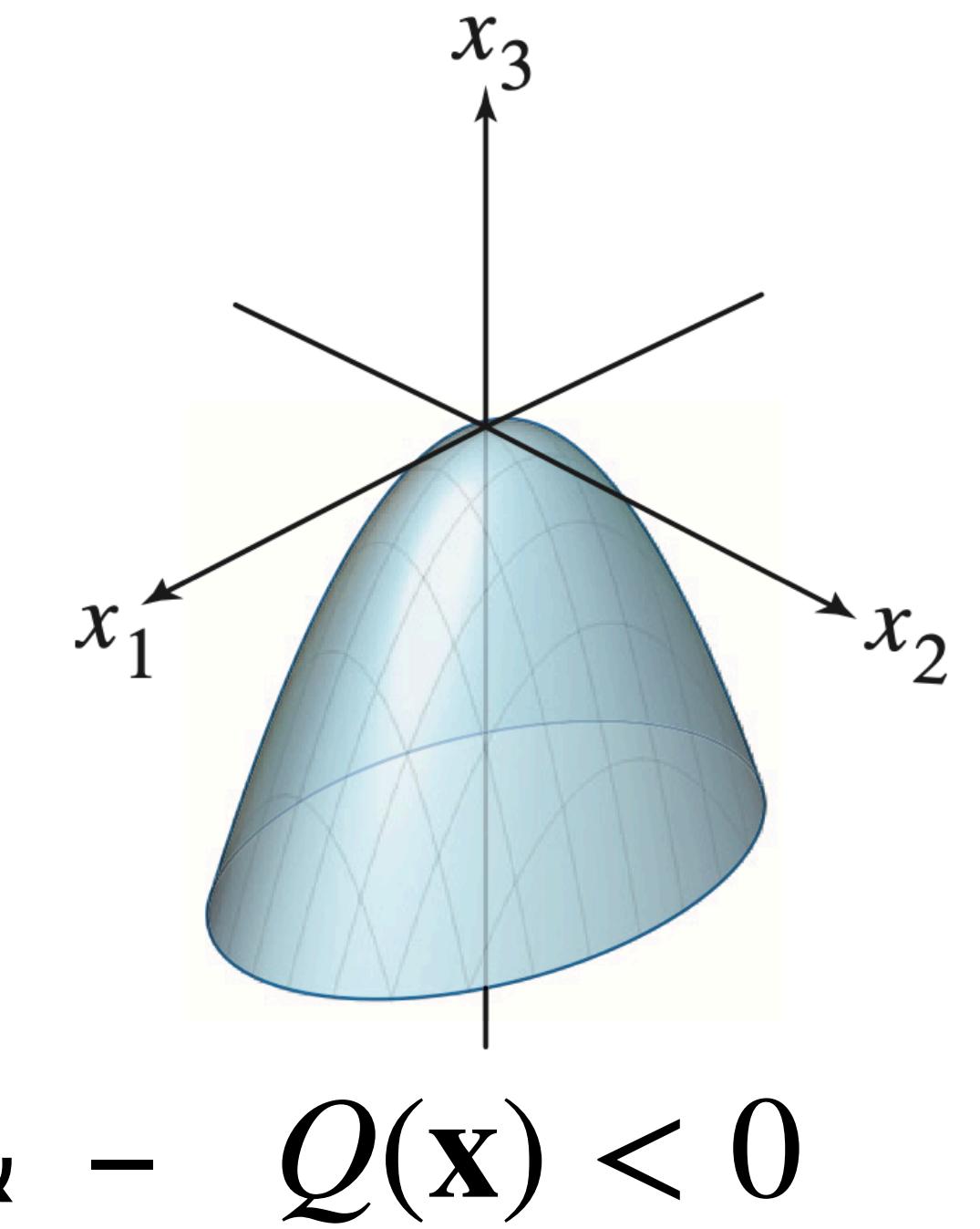
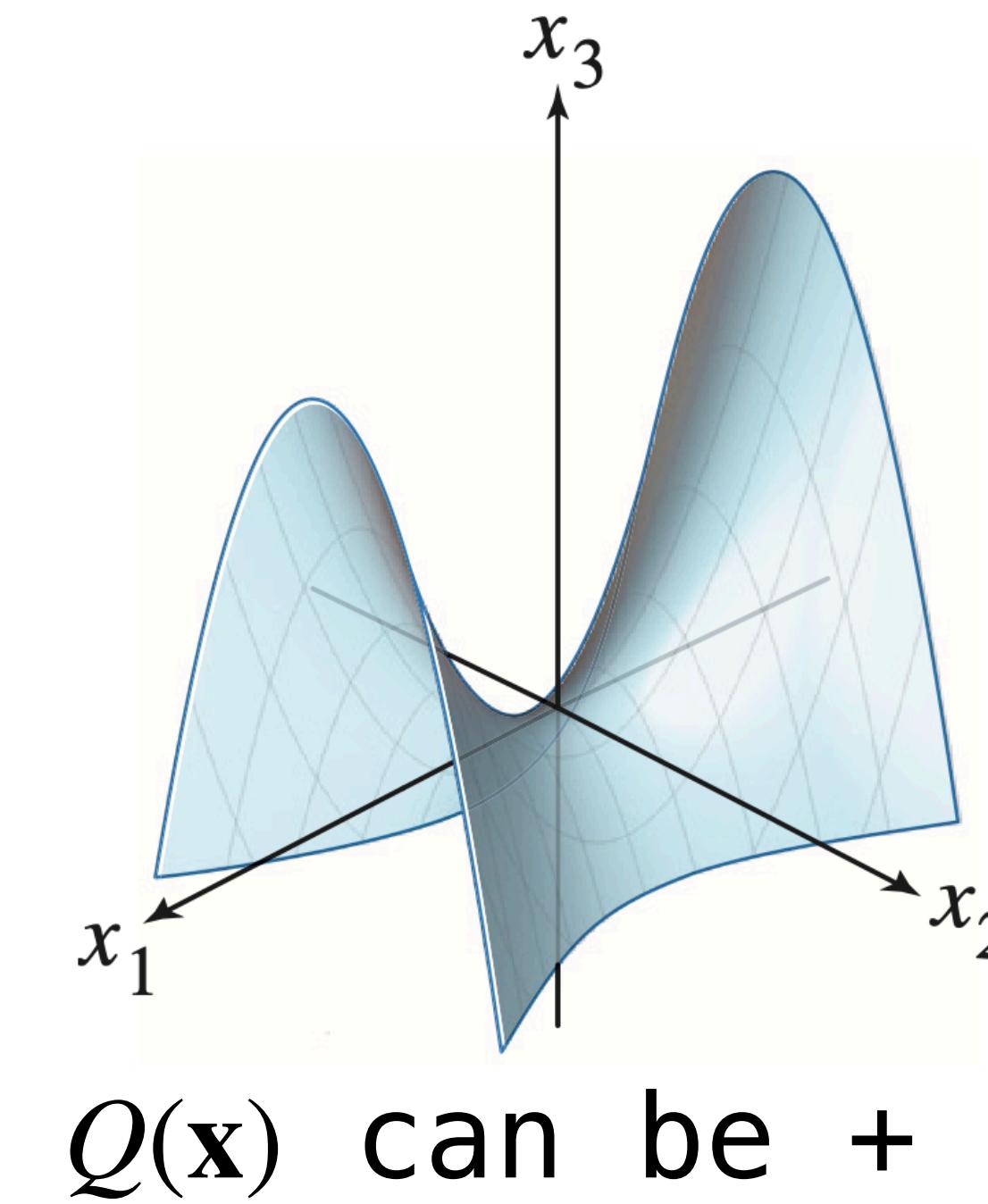
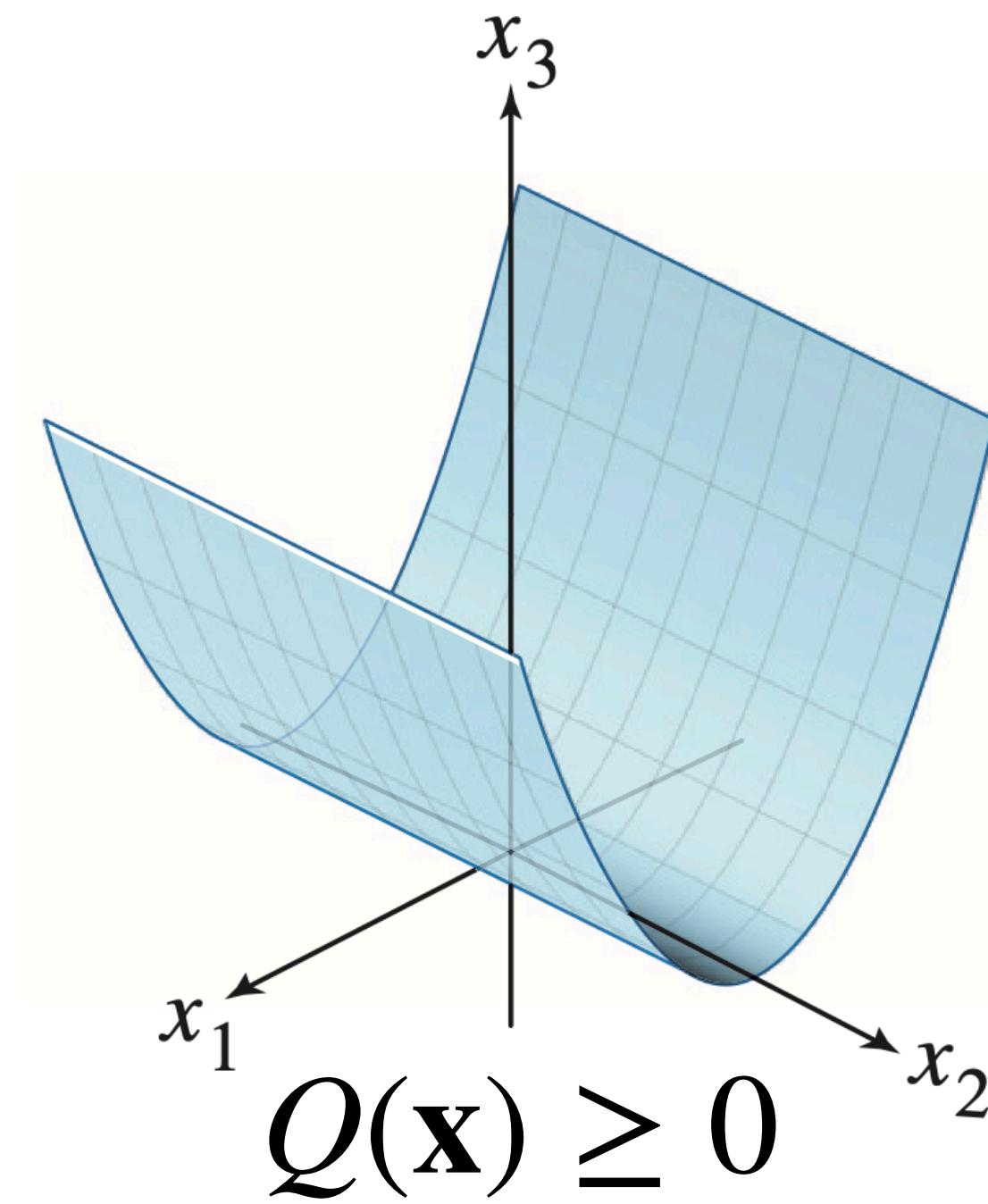
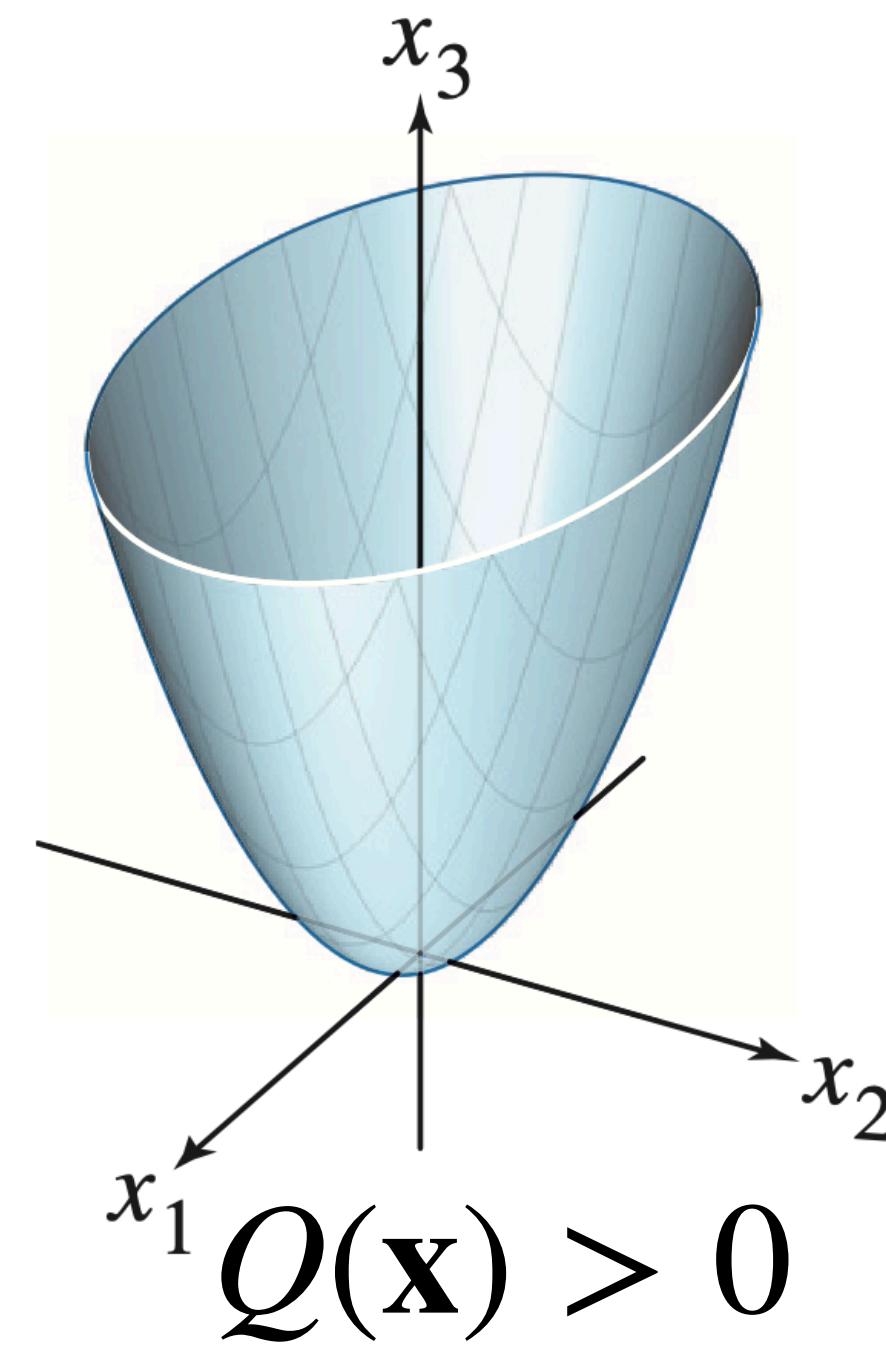
Shapes of Quadratic Forms



There are essentially three possible shapes (six if you include the negations).

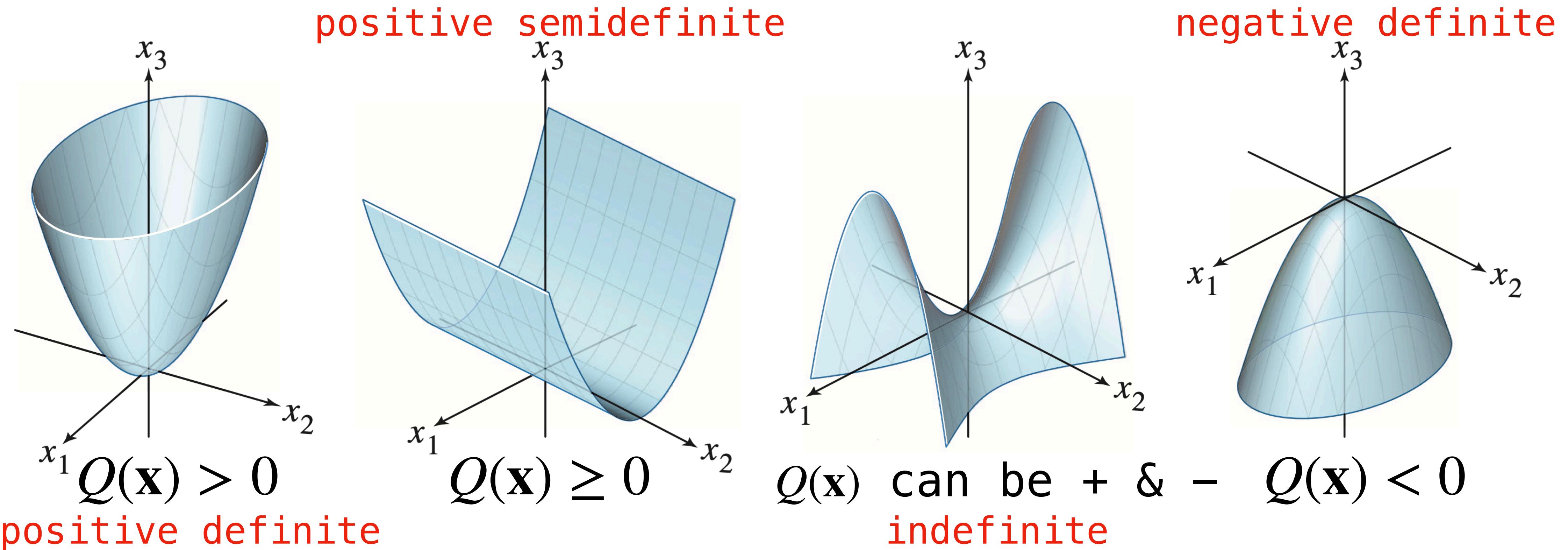
Can we determine what shape it will be mathematically?

Definiteness



For $\mathbf{x} \neq 0$, each of the above graphs satisfy the associated properties.

Definiteness



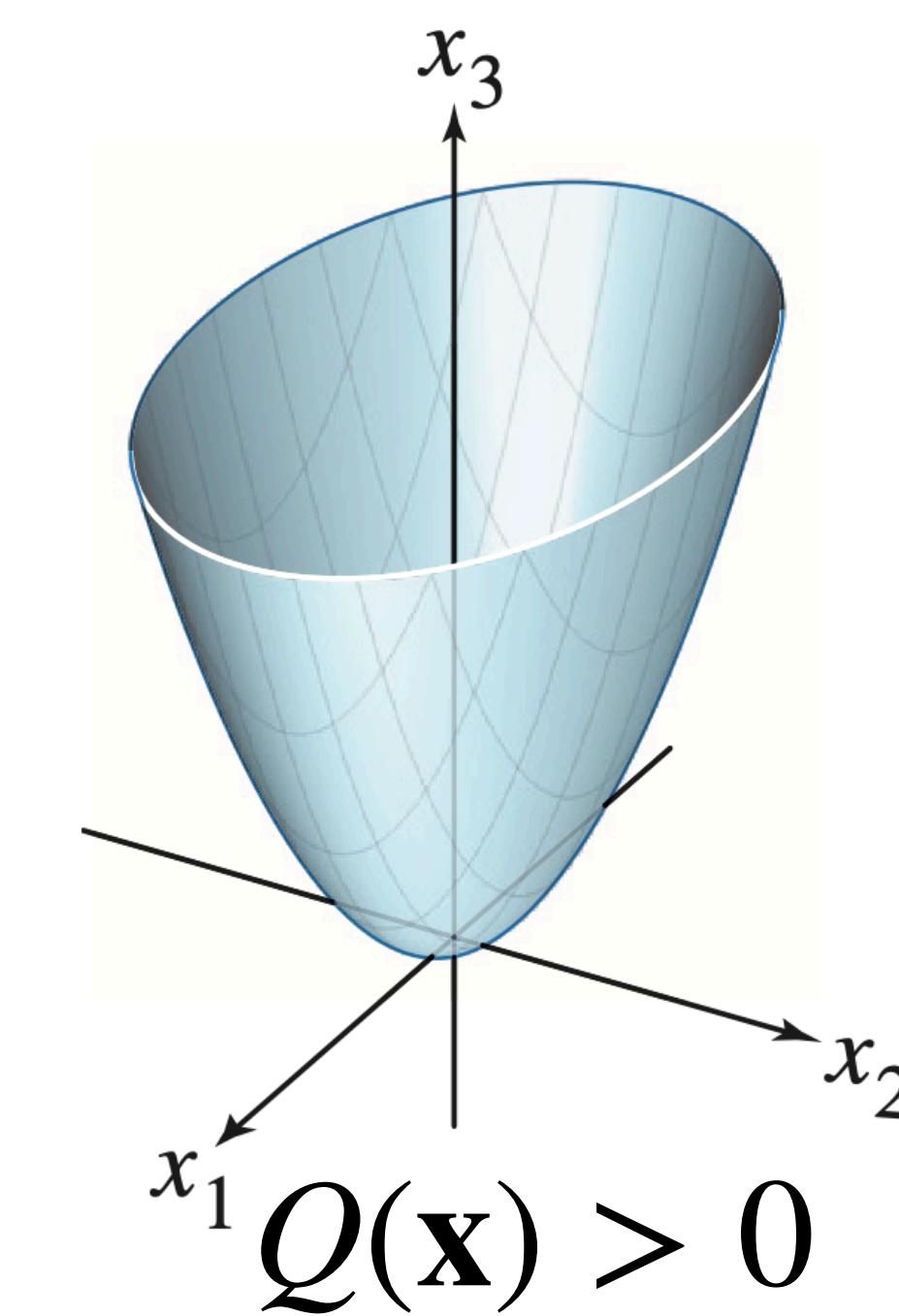
For $\mathbf{x} \neq 0$, each of the above graphs satisfy the associated properties.

Definiteness and Eigenvectors

Theorem. For a symmetric matrix A , the quadratic form $\mathbf{x}^T A \mathbf{x}$

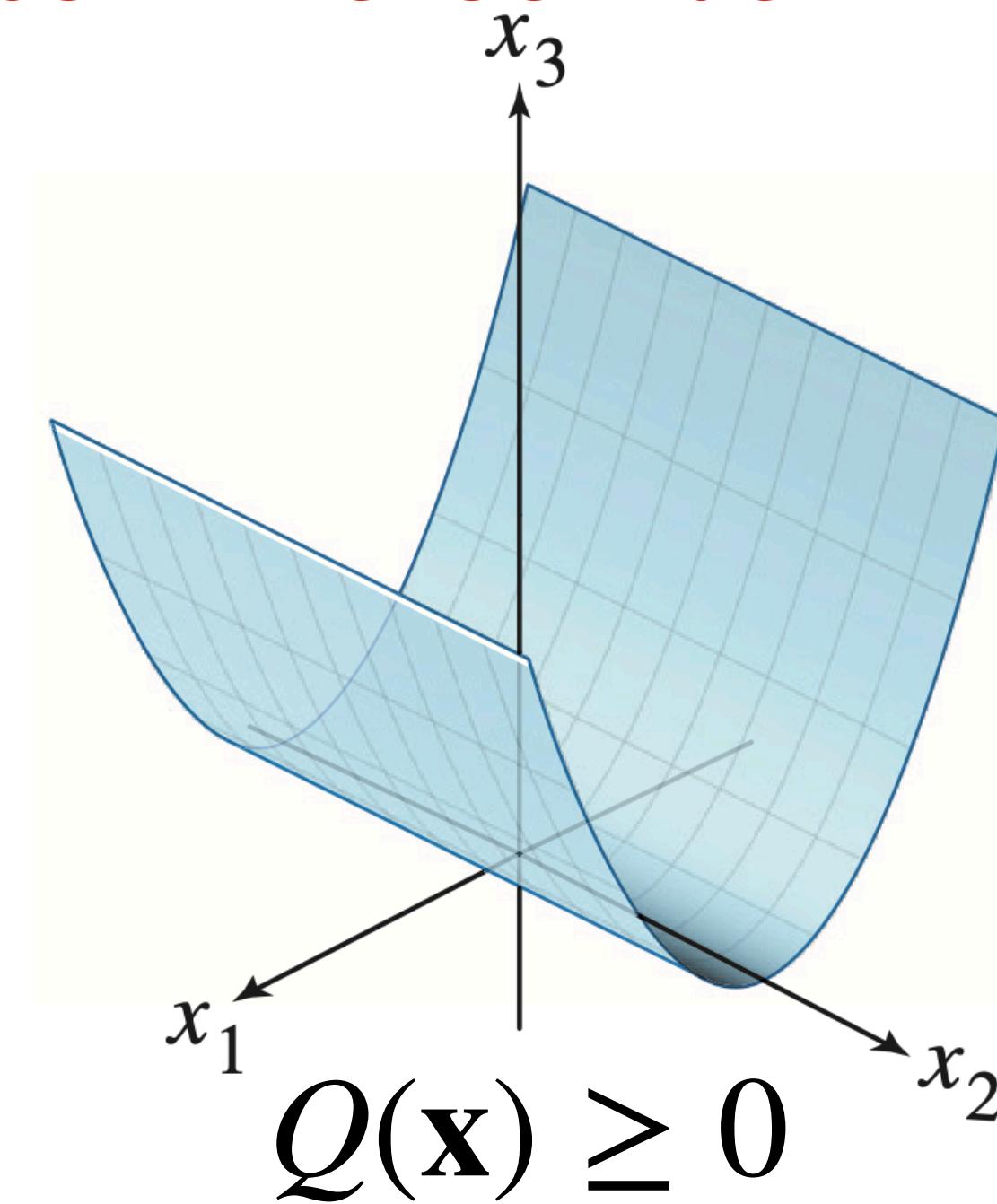
- » **positive definite** \equiv all positive eigenvalues
- » **positive semidefinite** \equiv all nonnegative eigenvalues
- » **indefinite** \equiv positive and negative eigenvalues
- » **negative definite** \equiv all negative eigenvalues

Definiteness

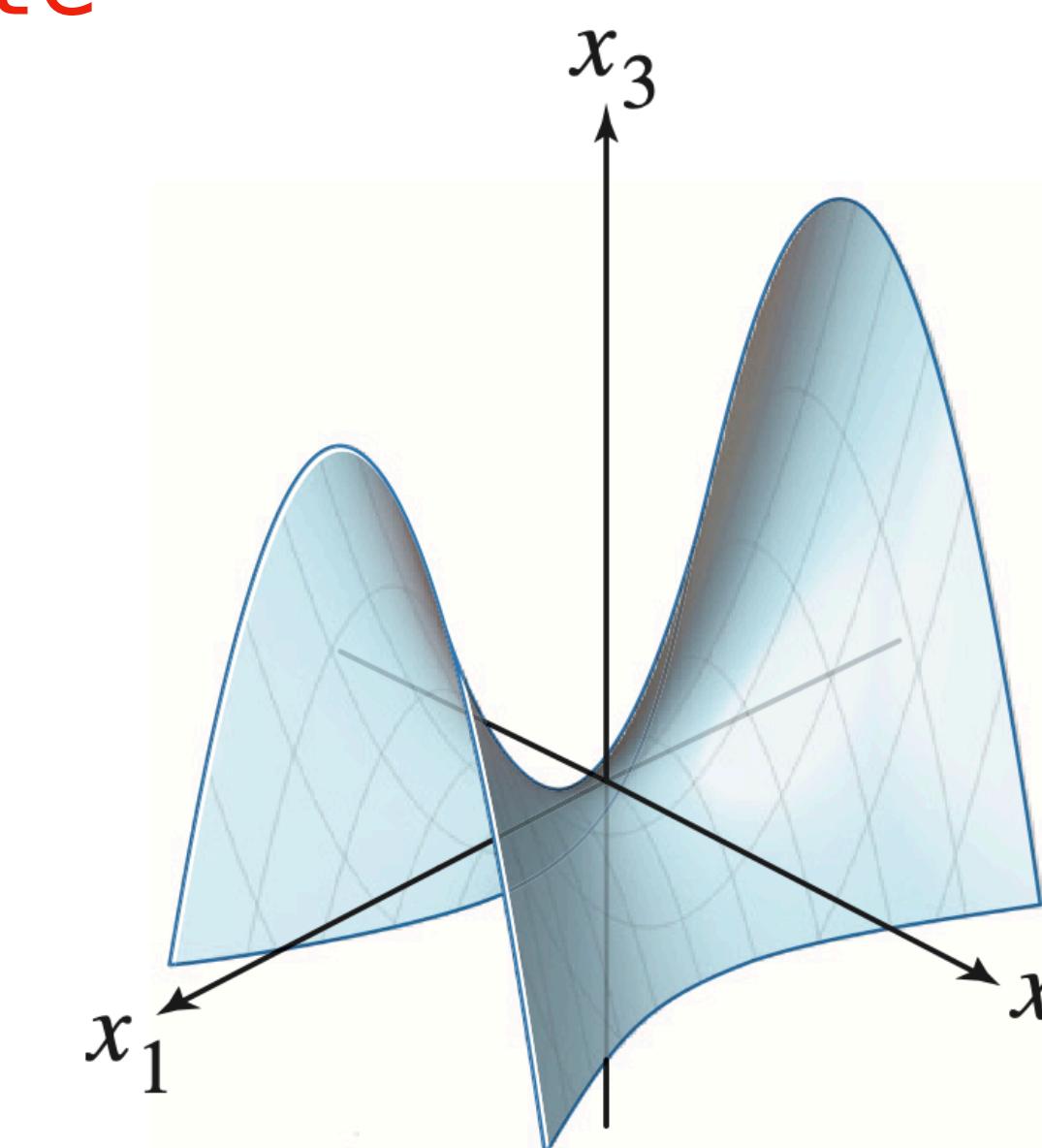


positive definite
all pos. eigenvals

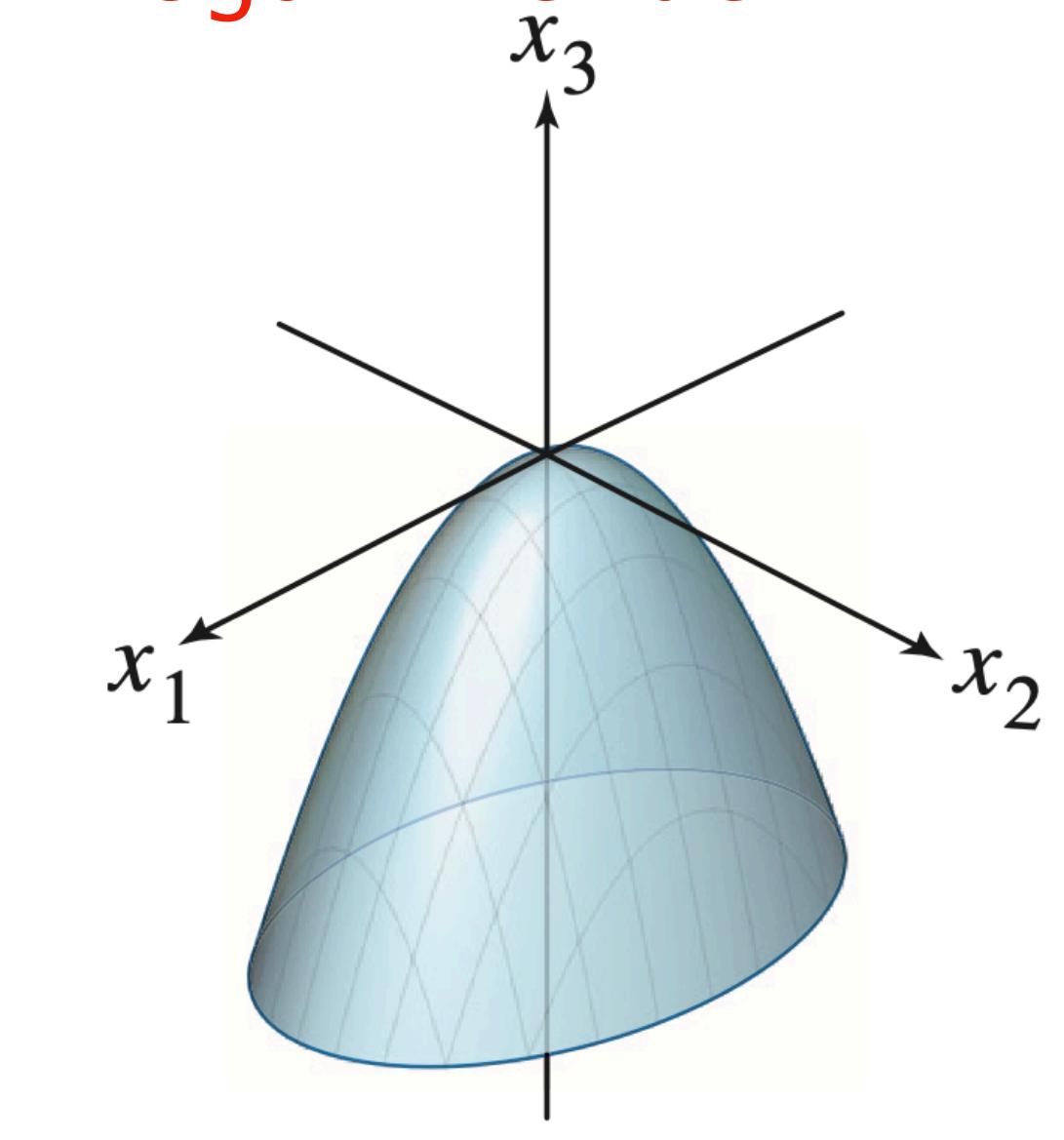
all nonneg. eigenvals
positive semidefinite



$Q(\mathbf{x})$ can be + & -
indefinite
pos. and neg. eigenvals



all neg. eigenvals
negative definite



Example

$$Q(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 4x_2x_3 + x_3^2$$

Let's determine which case this is:

Constrained Optimization

In General

In General

Given a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and a set of vectors X from \mathbb{R}^n the **constrained minimization problem** for f over X is the problem of determining

$$\min_{\mathbf{v} \in X} f(\mathbf{v})$$

In General

Given a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and a set of vectors X from \mathbb{R}^n the **constrained minimization problem** for f over X is the problem of determining

$$\min_{\mathbf{v} \in X} f(\mathbf{v})$$

(analogously for maximization)

In General

Given a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and a set of vectors X from \mathbb{R}^n the **constrained minimization problem** for f over X is the problem of determining

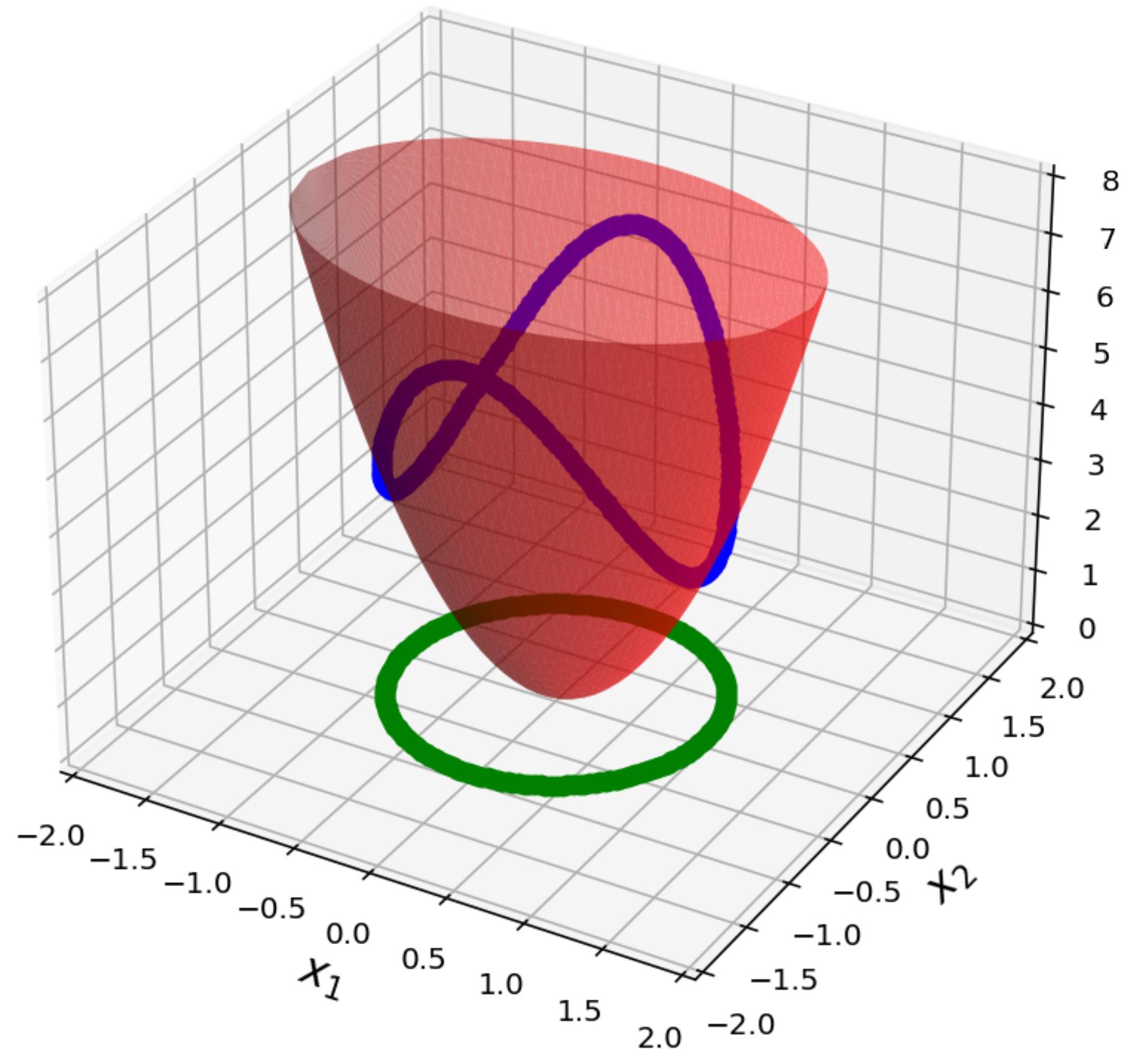
$$\min_{\mathbf{v} \in X} f(\mathbf{v})$$

(analogously for maximization)

Find the smallest value of $f(\mathbf{v})$ subject to choosing a vector in X

Constrained Optimization for Quadratic Forms and Unit Vectors

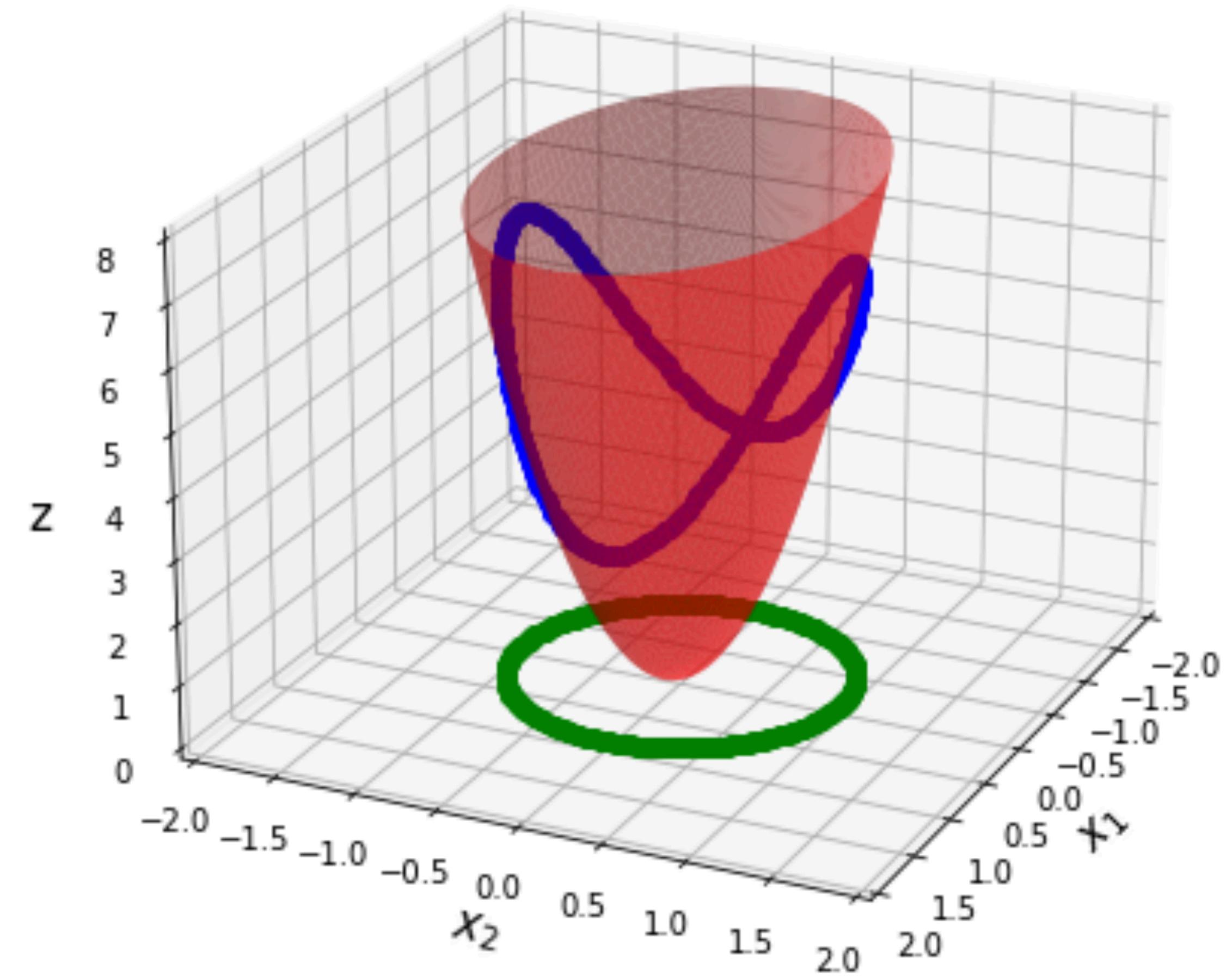
mini/maximize $\mathbf{x}^T \mathbf{A} \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$



It's common to constraint to unit vectors.

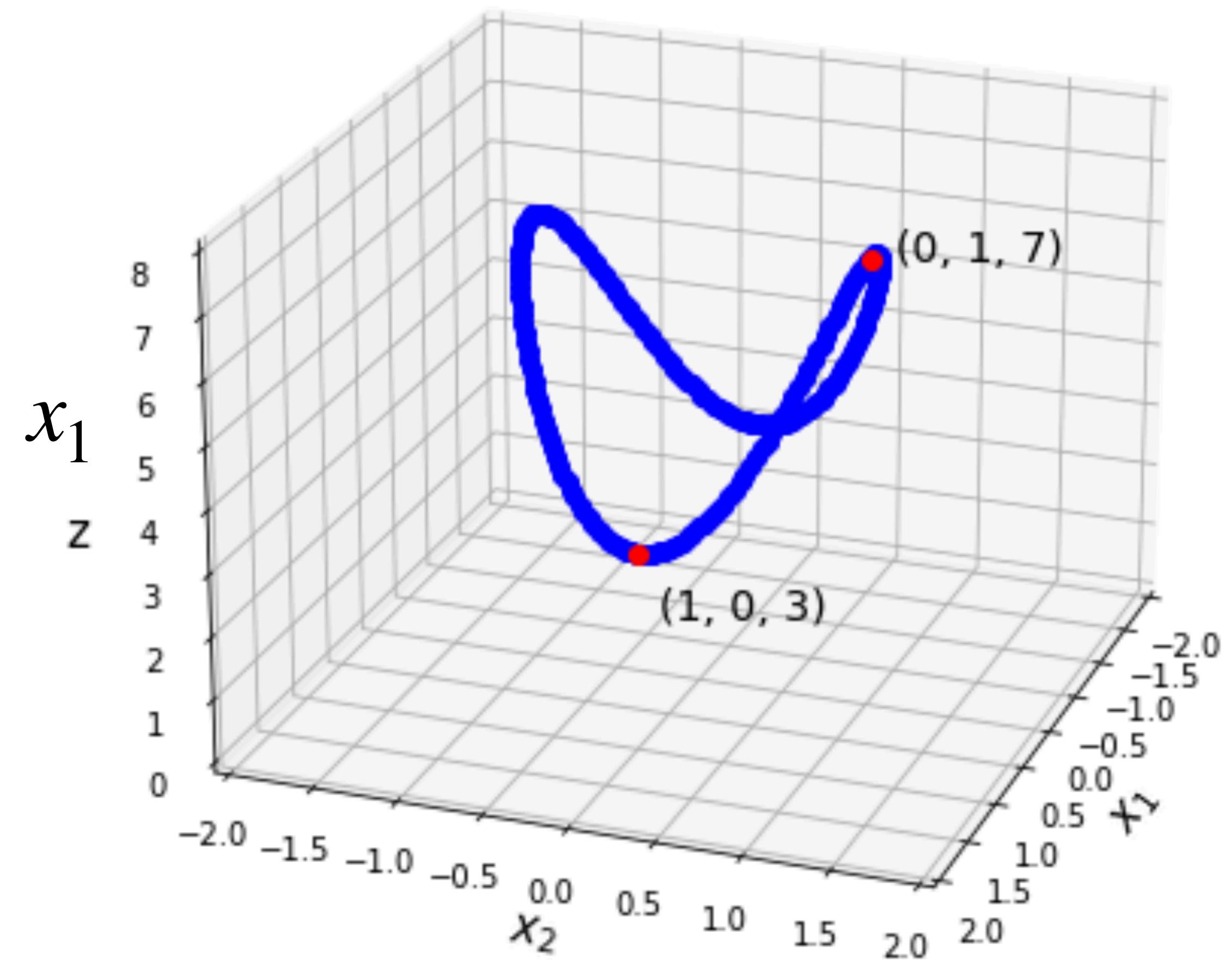
Example: $3x_1^2 + 7x_2^2$

What are the min/max values?:



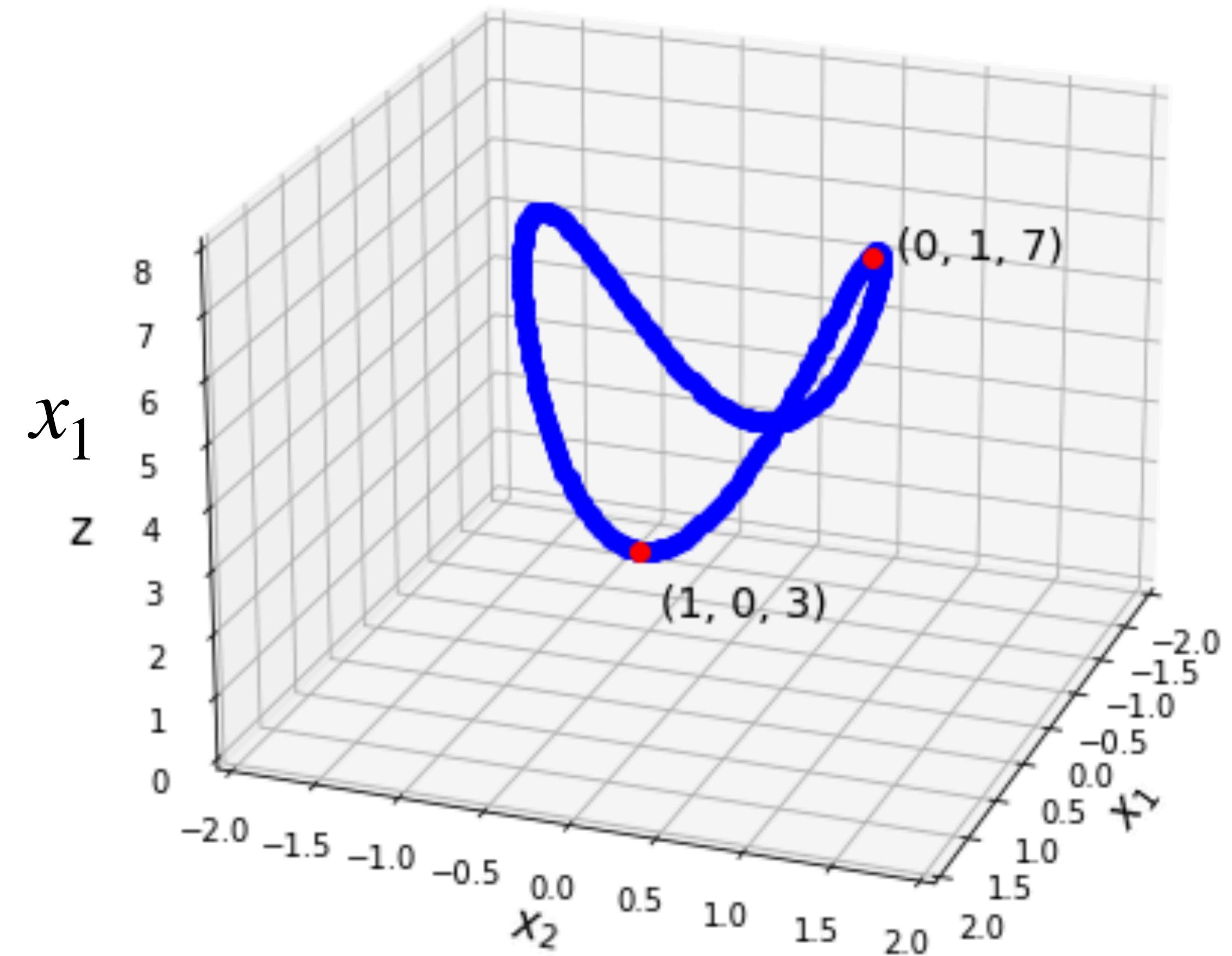
Example: $3x_1^2 + 7x_2^2$

The minimum and maximum values are attained when the "weight" of the vector is distributed all on x_1 or x_2 .



Example: $3x_1^2 + 7x_2^2$

What is the matrix?:



Constrained Optimization and Eigenvalues

Theorem. For a symmetric matrix A , with largest eigenvalue λ_1 and smallest eigenvalue λ_n

$$\max_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = \lambda_1$$

$$\min_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = \lambda_n$$

argmax is \vec{v}_1 eigenvector

argmin is \vec{v}_n eigenvector

No matter the shape of A , this will hold.

How To: Constrained Optimization

How To: Constrained Optimization

Problem. Find the maximum value of $\mathbf{x}^T A \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$.

How To: Constrained Optimization

Problem. Find the maximum value of $\mathbf{x}^T A \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$.

Solution. Find the largest eigenvalue of A , this will be the maximum value.

How To: Constrained Optimization

Problem. Find the maximum value of $\mathbf{x}^T A \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$.

Solution. Find the largest eigenvalue of A , this will be the maximum value.

(Use NumPy)

Practice Problem

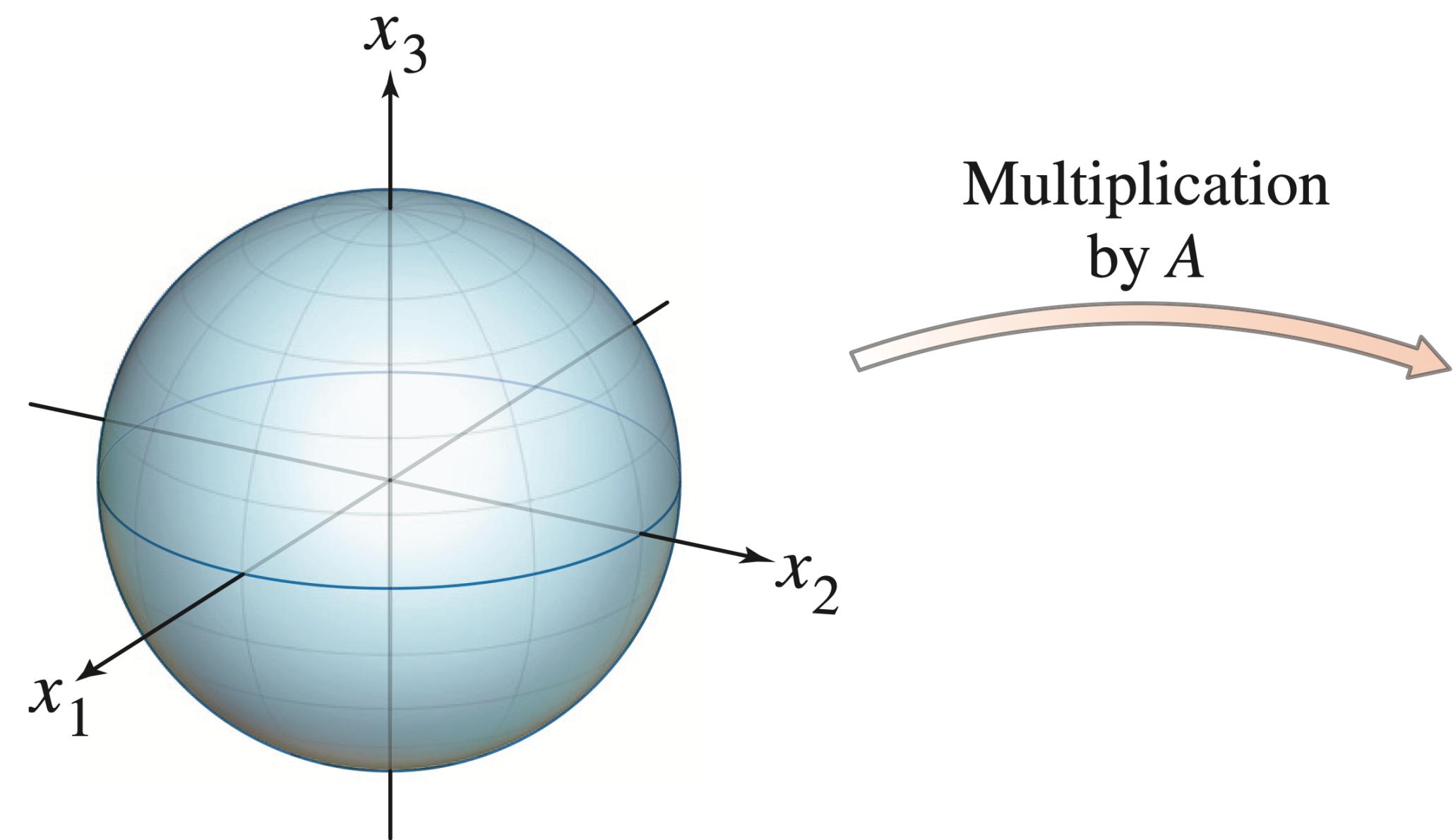
$$Q(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 4x_2x_3 + x_3^2$$

Find the maximum value of $Q(\mathbf{x})$ subject to $\|\mathbf{x}\| = 1$

Singular Value Decomposition (Looking Ahead)

Question

What shape is a the unit sphere after a linear transformation?

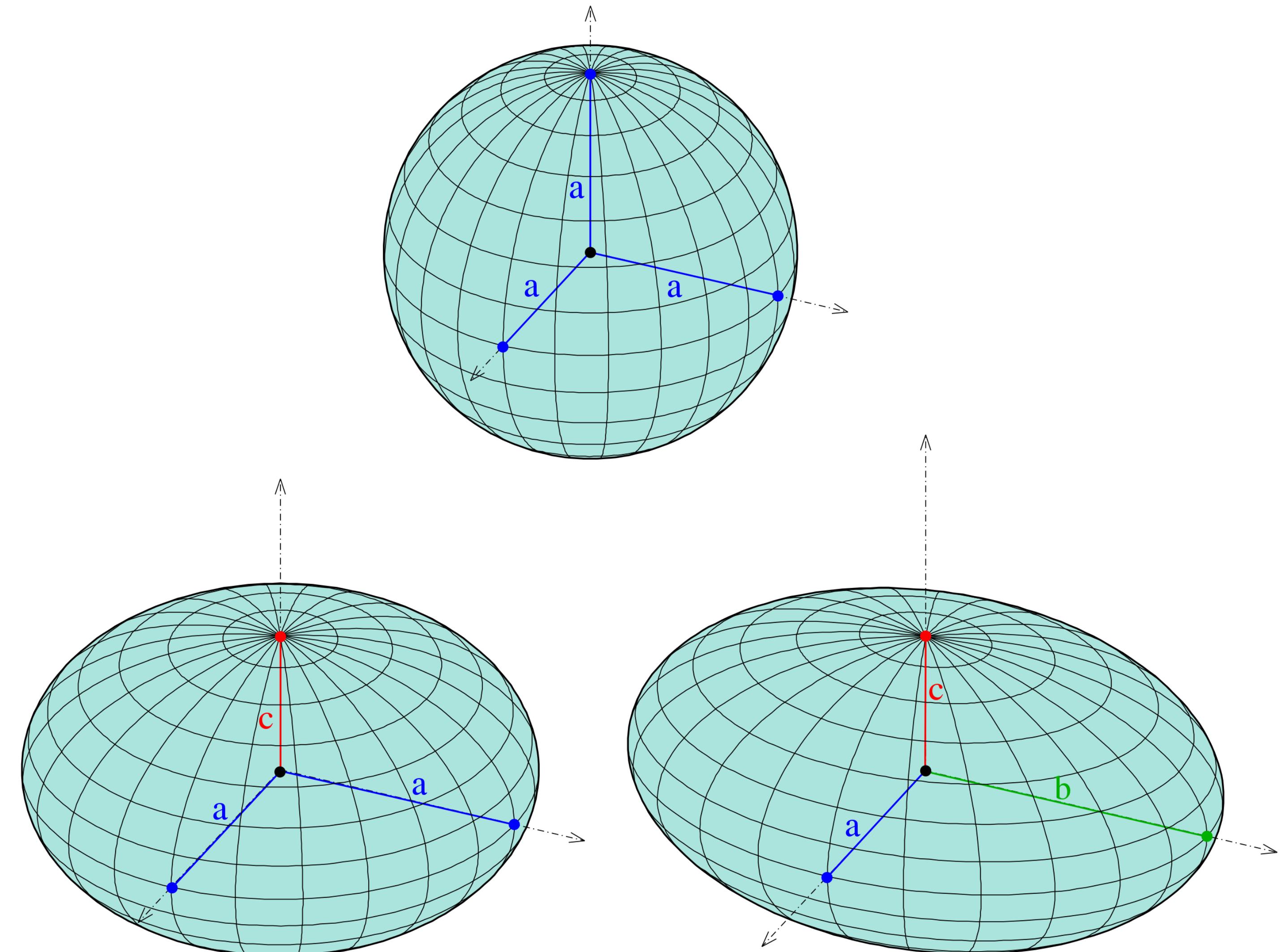


???

Ellipsoids

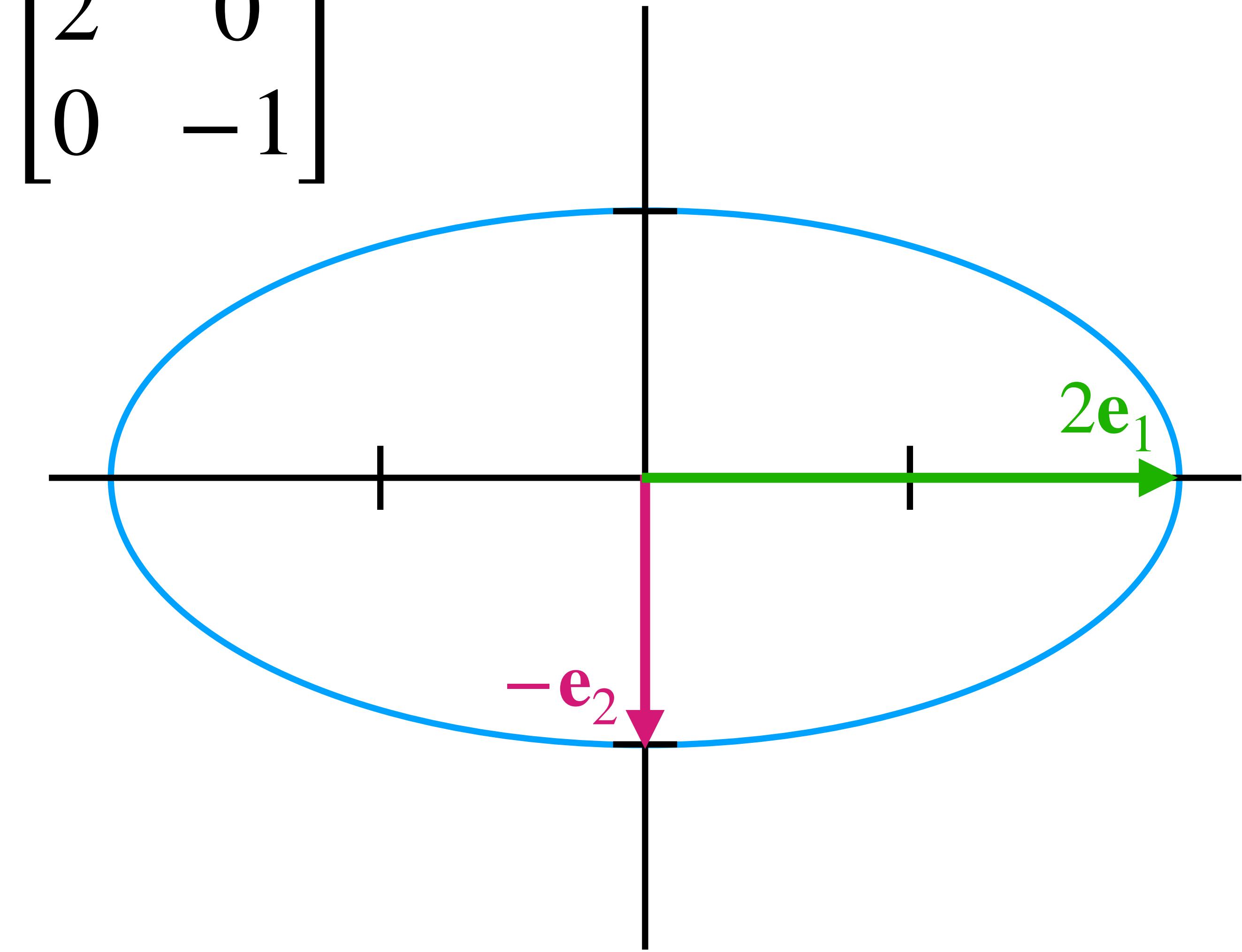
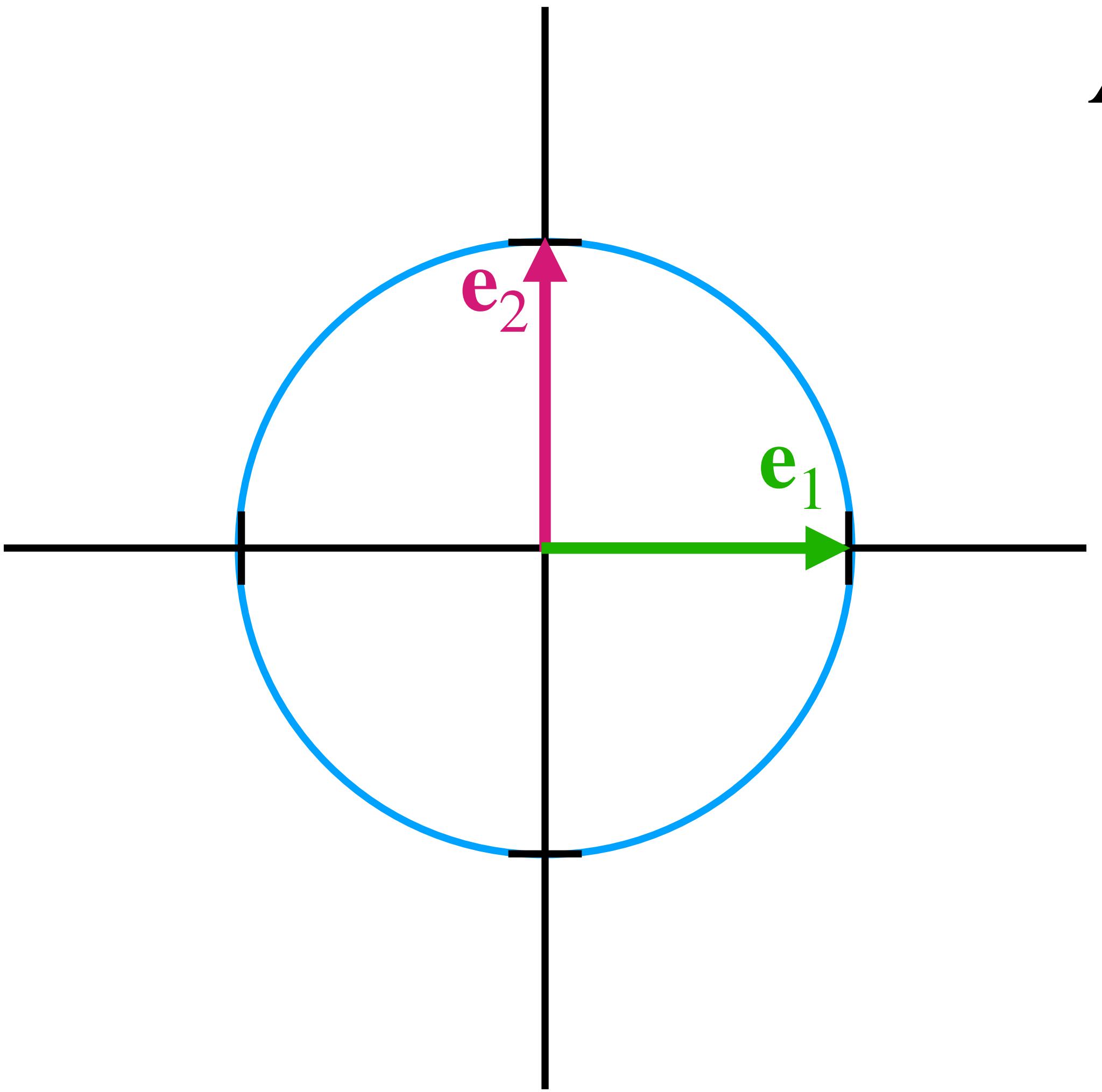
Ellipsoids are spheres "stretched" in orthogonal directions called the **axes of symmetry** or the **principle axes**.

Linear transformations maps spheres to ellipsoids.

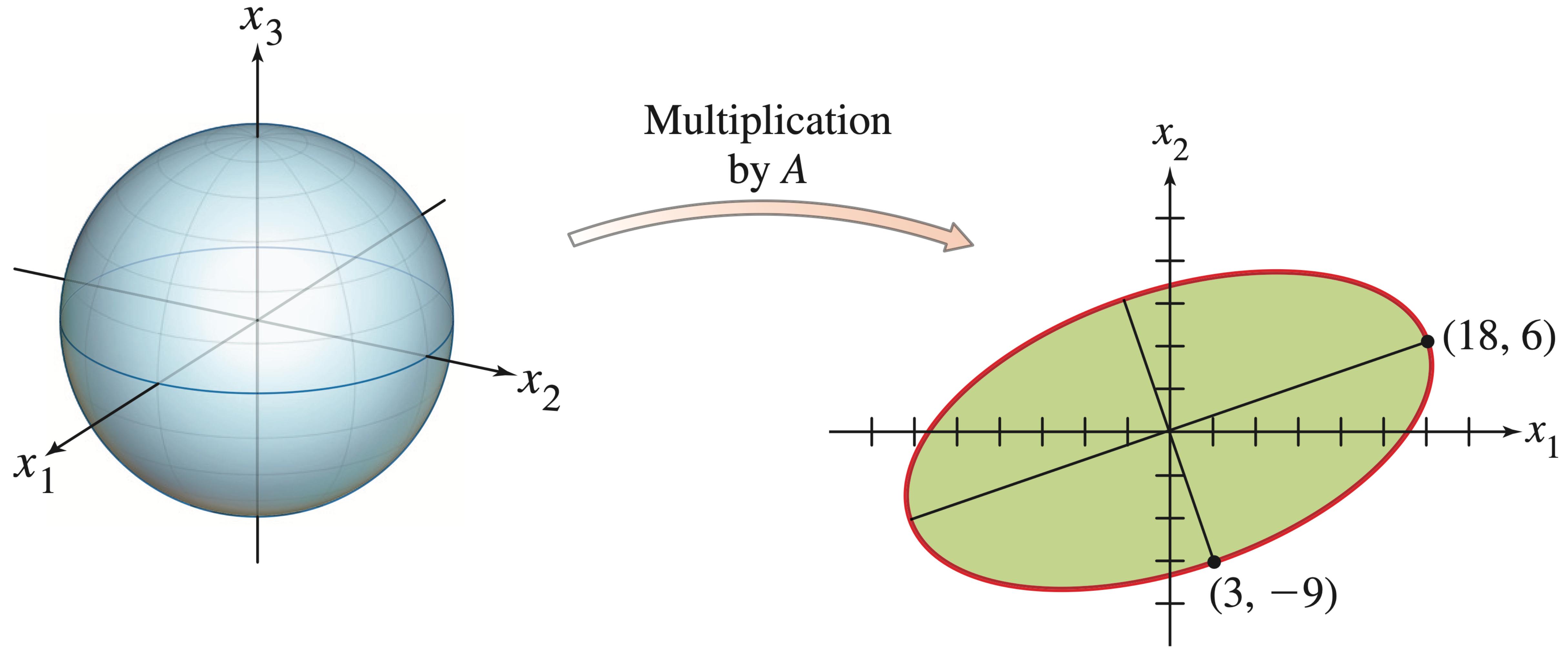


Simple Example : Scaling Matrices

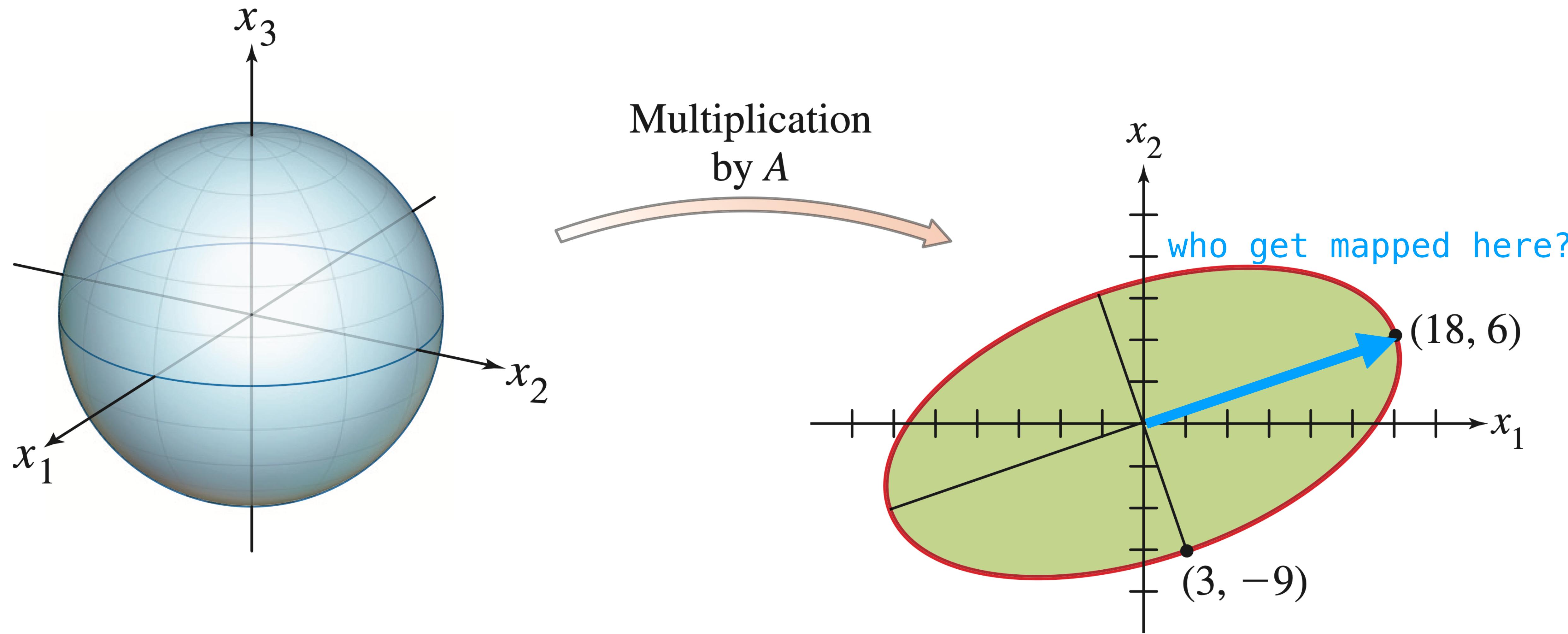
$$A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$



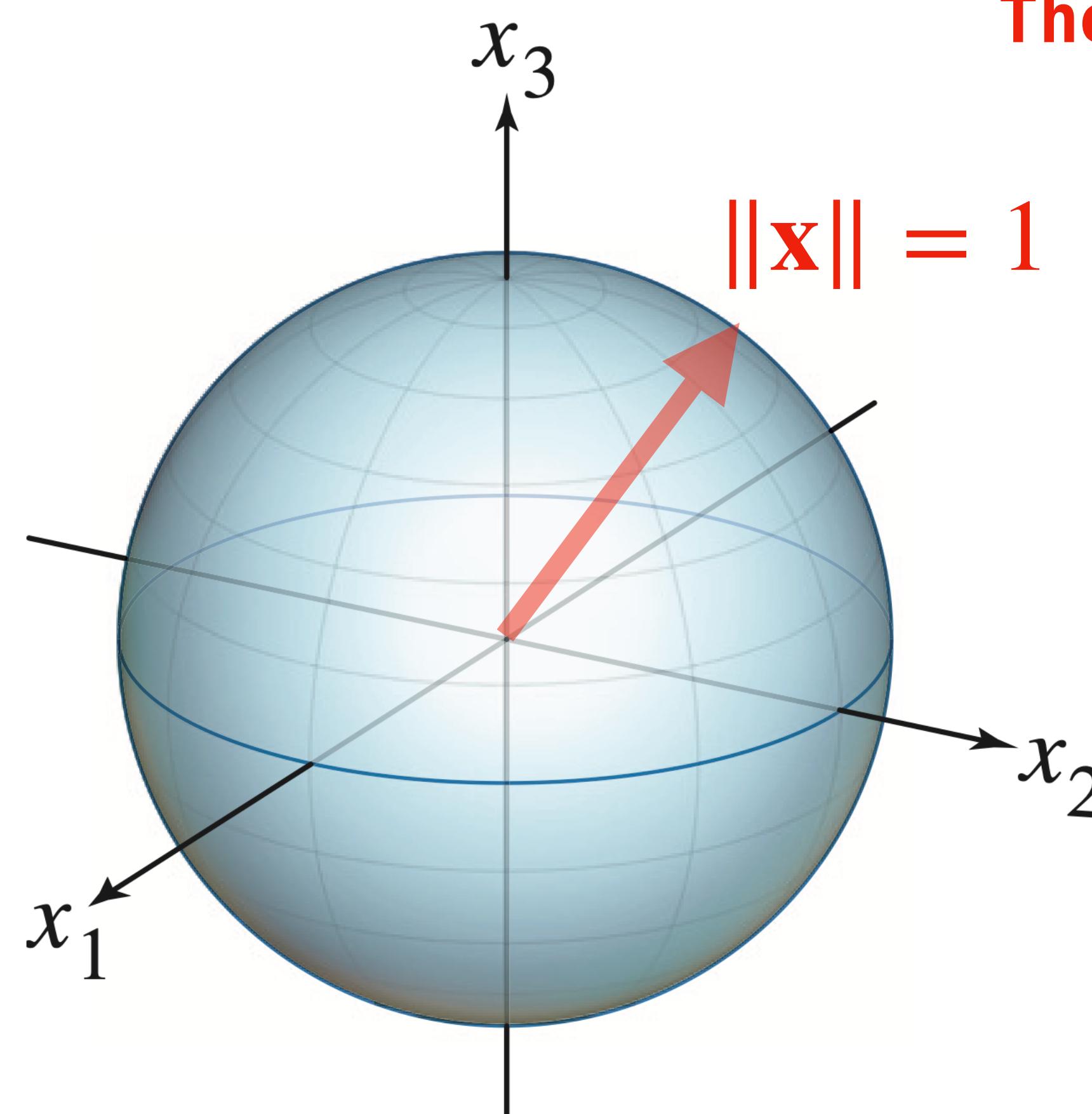
The Picture



The Picture

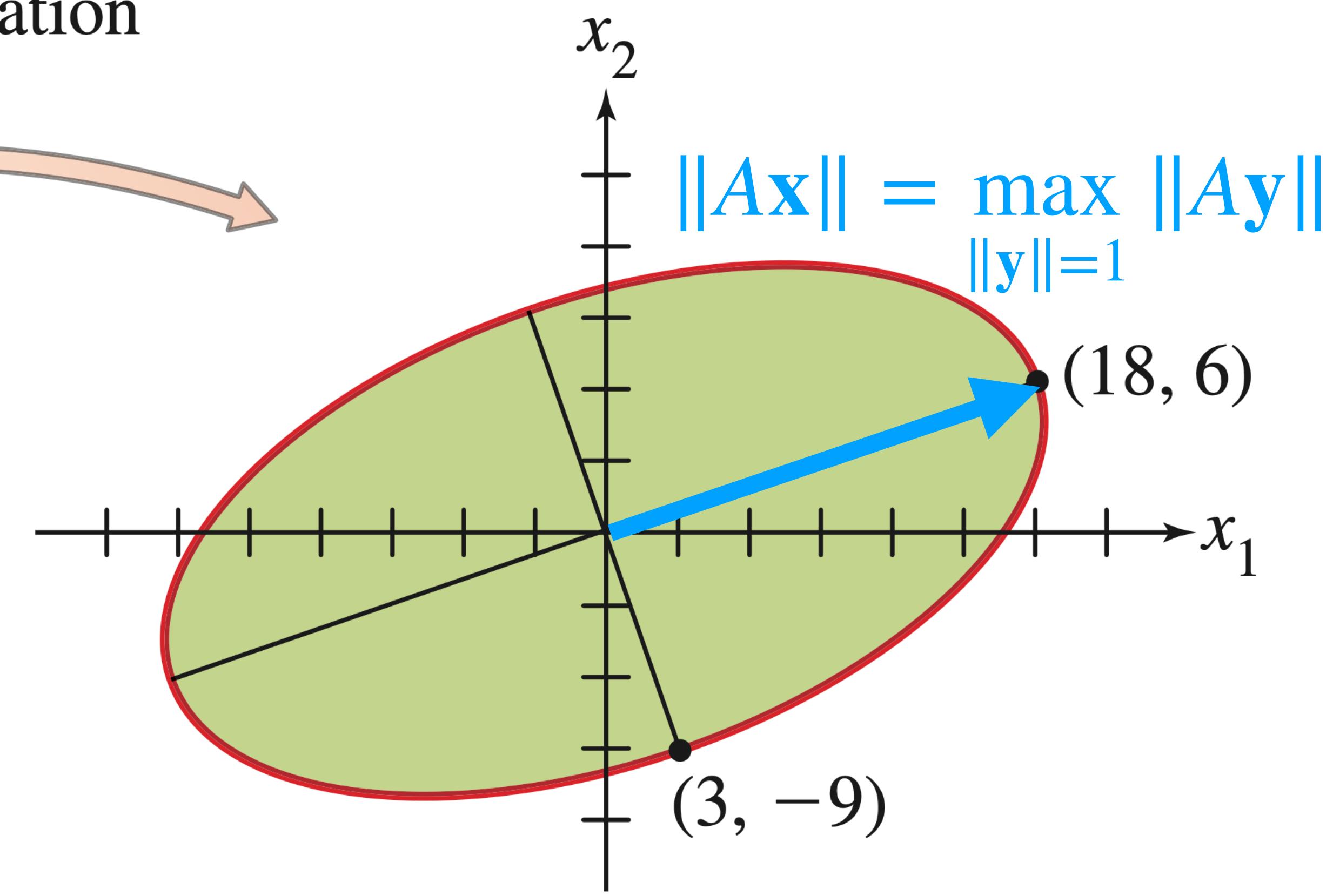


The Picture

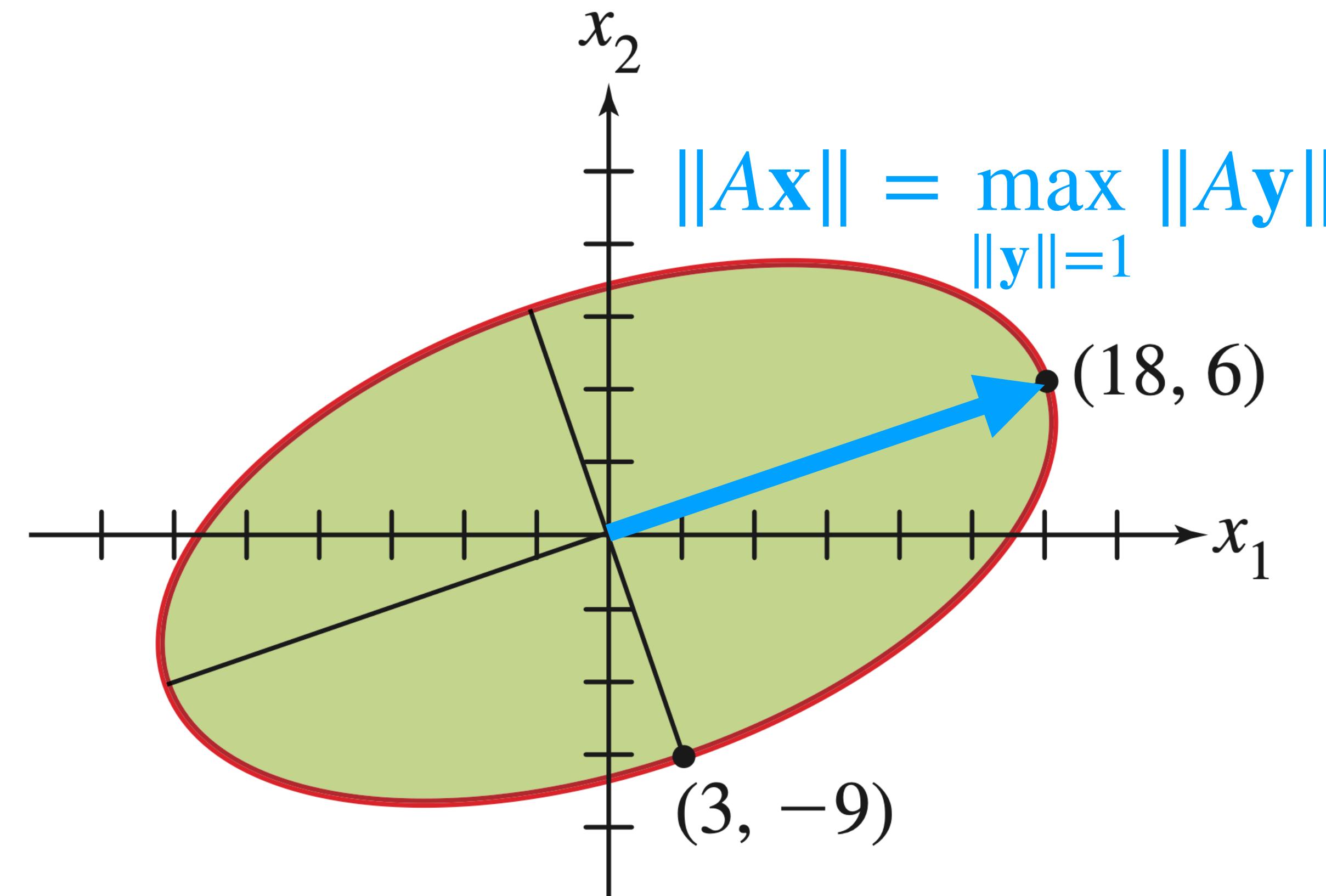


The longest end of the ellipse is the solution to a constrained optimization problem

Multiplication
by A

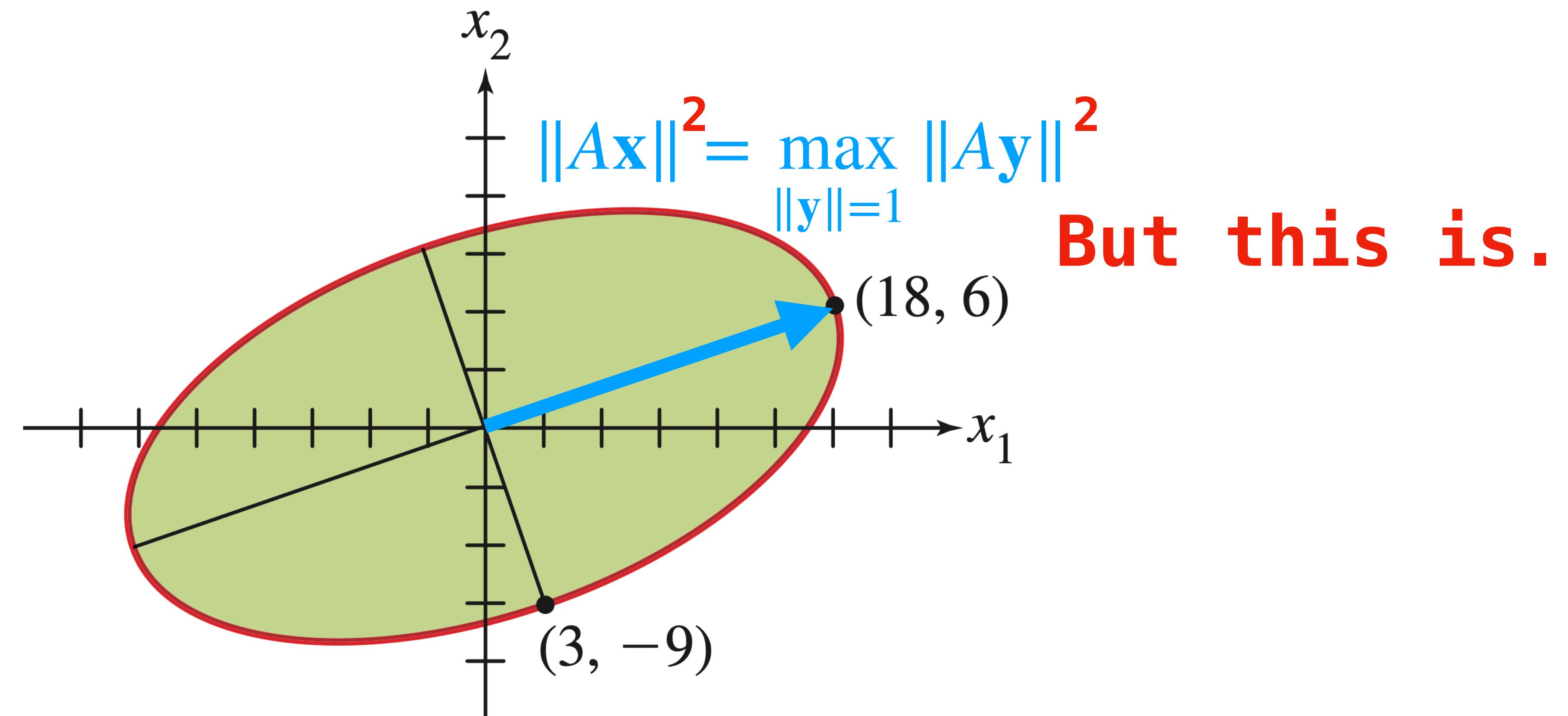


The Picture



This is not a quadratic form...

The Picture

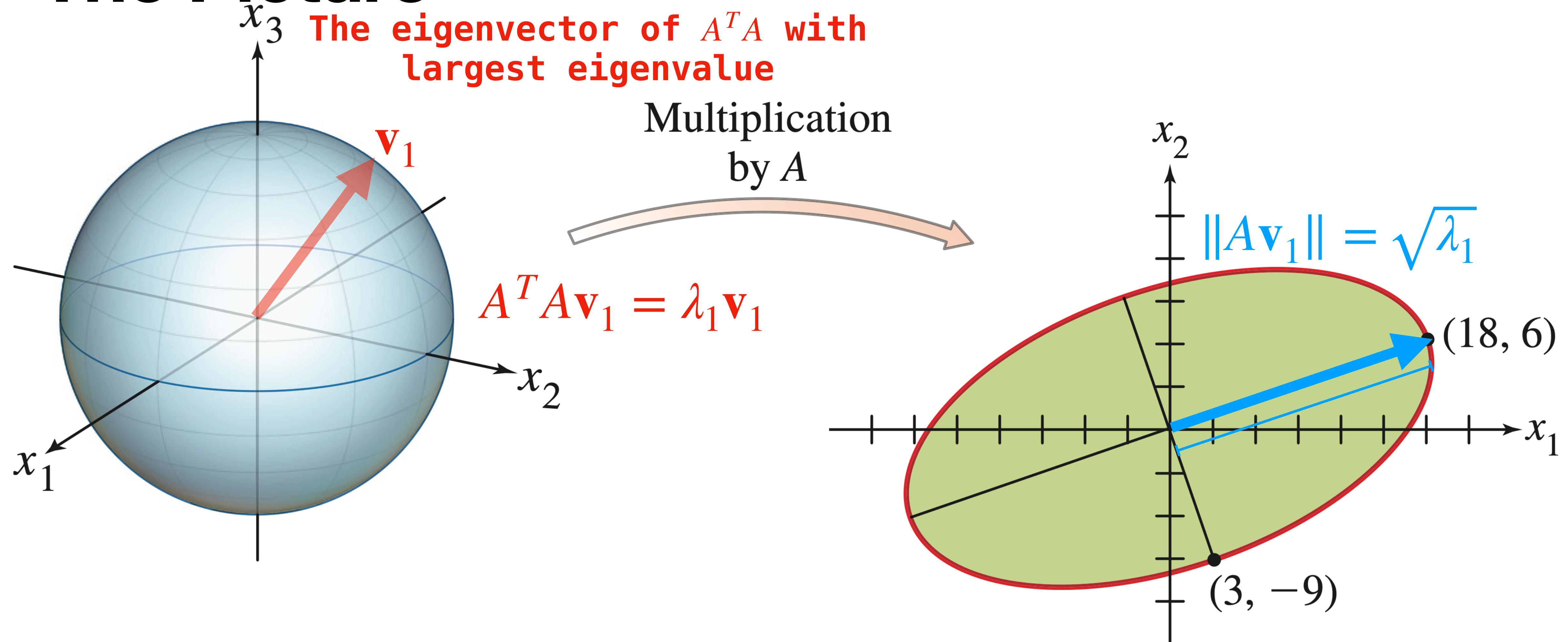


This is not a quadratic form...

A Quadratic Form

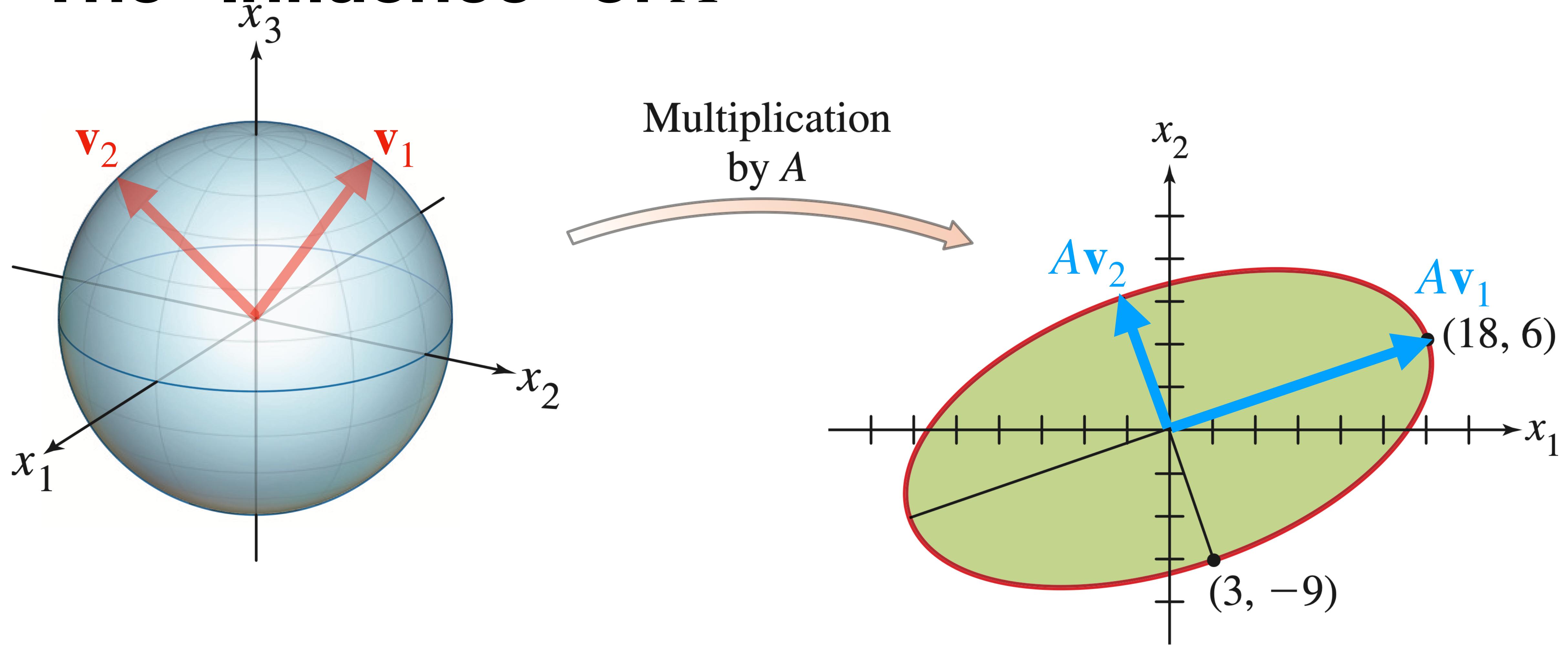
What does $\|Ax\|^2$ look like?:

The Picture



\mathbf{v}_1 solves the constrained optimization problem.

The "Influence" of A



v_1 is "most affected" by A and v_2 is "least affected"

Properties of $A^T A$

Properties of $A^T A$

- » It's symmetric.

Properties of $A^T A$

- » It's symmetric.
- » So its orthogonally diagonalizable.

Properties of $A^T A$

- » It's symmetric.
- » So its orthogonally diagonalizable.
- » There is an orthogonal basis of eigenvectors.

Properties of $A^T A$

- » It's symmetric.
- » So its orthogonally diagonalizable.
- » There is an orthogonal basis of eigenvectors.
- » Its eigenvalues are nonnegative.

Properties of $A^T A$

- » It's symmetric.
- » So its orthogonally diagonalizable.
- » There is an orthogonal basis of eigenvectors.
- » Its eigenvalues are nonnegative.
- » It's positive semidefinite.

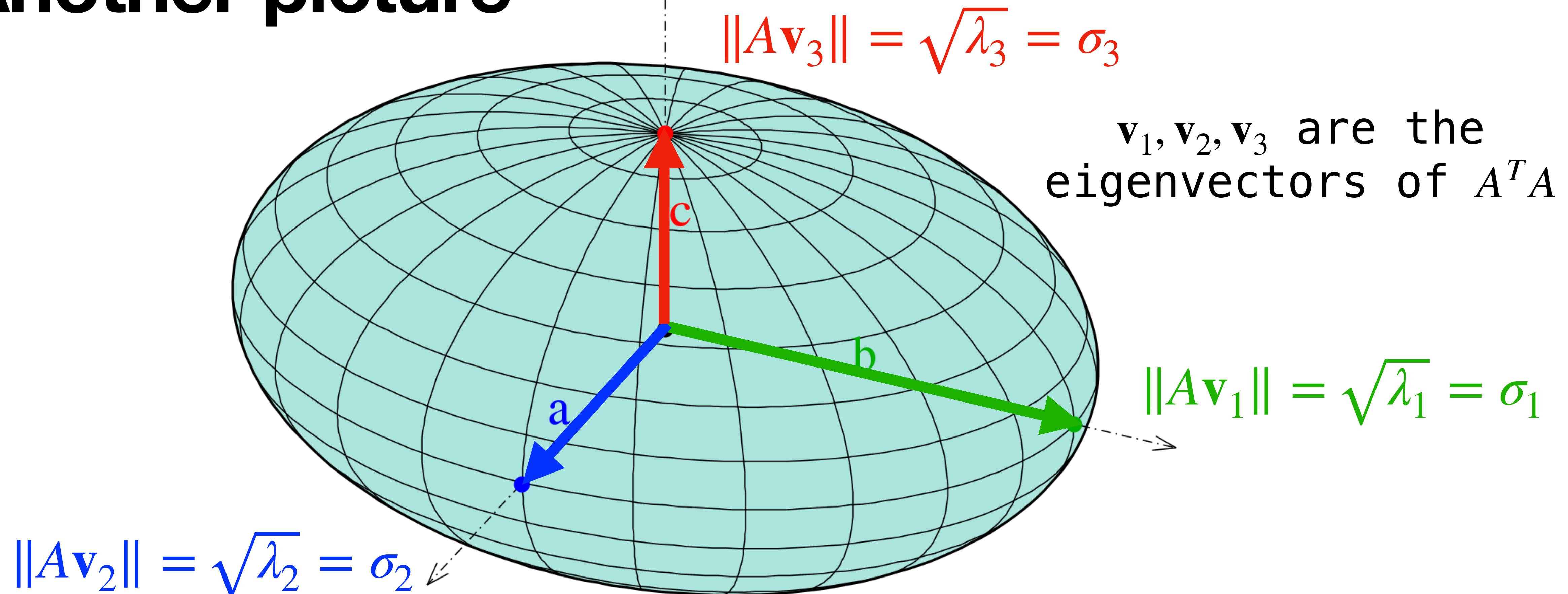
Singular Values

Definition. For an $m \times n$ matrix A , the **singular values** of A are the n values

$$\sigma_1 \geq \sigma_2 \dots \geq \sigma_n \geq 0$$

where $\sigma_i = \sqrt{\lambda_i}$ and λ_i is an eigenvalue of $A^T A$.

Another picture



The **singular values** are the lengths of the axes of *symmetry* of the ellipsoid after transforming the unit sphere.

Every $m \times n$ matrix transforms the unit m -sphere into an n -ellipsoid.

So every $m \times n$ matrix has
 n singular values.