Symmetric Matrices

Geometric Algorithms
Lecture 25

CAS CS 132

Recap Problem

$X\vec{\beta} = \vec{y}$ $\quad X^T X \hat{\beta} = X^T \vec{y}$

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

Find the matrices X ← design matrix *as in the previous example to find the least squares best fix parabola <u>and the least squares best fit cubic</u> for this dataset.*

$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 \Rightarrow \beta_0 + \beta_1(0) + \beta_2(0)^2 = 3$

$\beta_0 + \beta_1(1) + \beta_2(1)^2 = 1$

⋮ (2 more)

"X" → design matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

$A\vec{x} = \vec{b}$

$A^T A \hat{x} = A^T \vec{b}$

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

Answer

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

↑ design matrix X

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 $A^T = A$.

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

$$A_{ij} = A_{ji} \quad \forall \text{ indices } i,j$$

Orthogonal Eigenvectors

If λ is an eigenvalue, Null(A) is corresponding eigenspace

$A\vec{v} = 0\vec{v} = 0$

$A\vec{v} = \lambda\vec{v}$ 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.

λ_1 & λ_2, respectively

$\lambda_1 \neq \lambda_2$

Verify: $\vec{u}^T A \vec{v} = \vec{u}^T \lambda_2 \vec{v} = \lambda_2 \vec{u}^T \vec{v}$

$\|$

$\vec{u}^T A^T \vec{v} = (A\vec{u})^T \vec{v} = (\lambda_1 \vec{u})^T \vec{v} = \lambda_1 \vec{u}^T \vec{v}$

$(\lambda_1 - \lambda_2)\vec{u}^T\vec{v} = 0 \Rightarrow \vec{u}^T\vec{v} = 0$

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 $A = PDP^{-1}$.

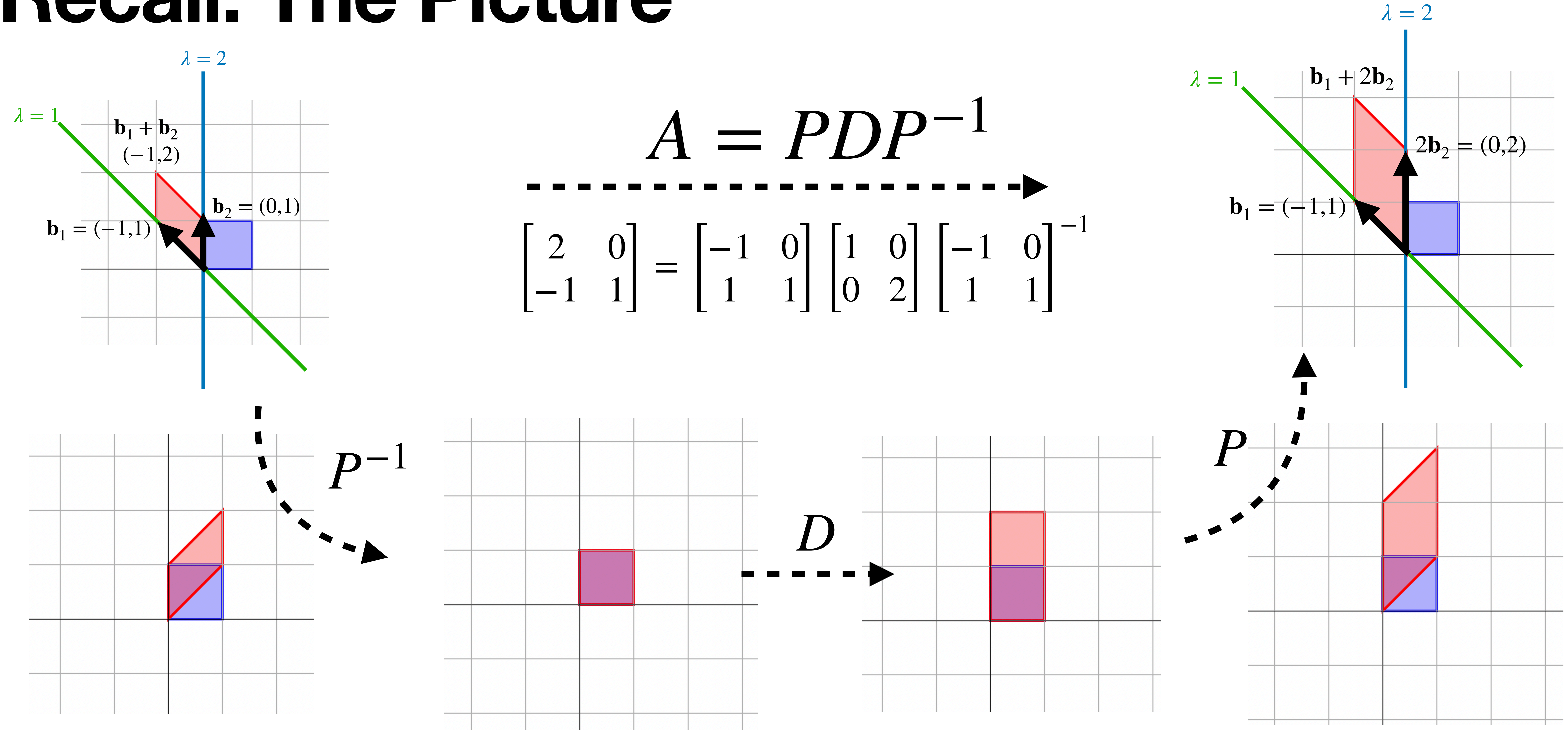
Recall: Diagonalizable Matrices

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



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

eigenbasis

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

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

$$A = \underbrace{P}_{\text{eigenbasis}}\underbrace{D}_{\text{eigenvalues}}P^{-1}$$

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.

Recall: The Diagonalization Theorem

eigenbasis

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

eigenvalues

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

(not all matrices are diagonalizable $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$)

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.

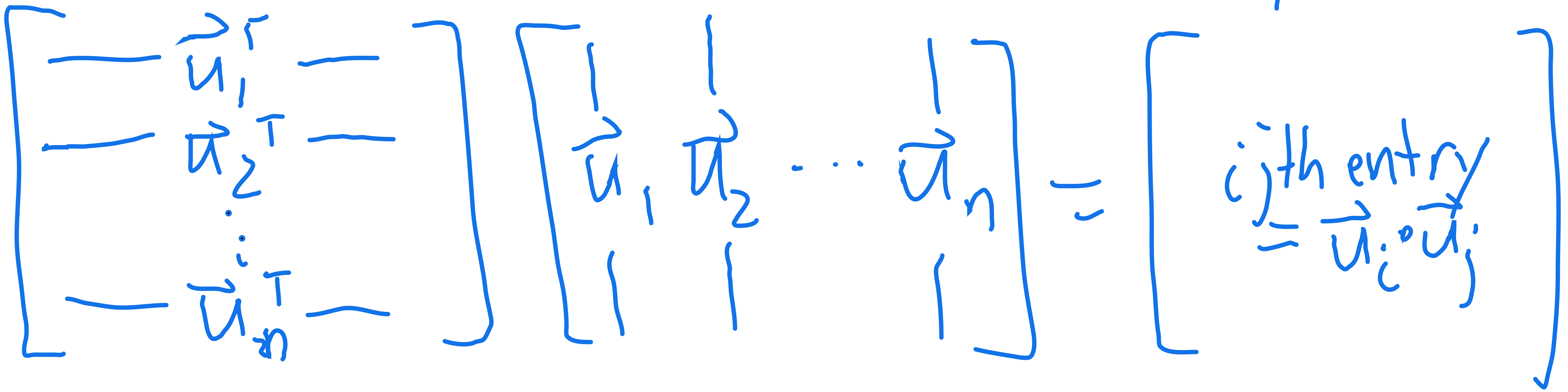
The notes call a square orthonormal matrix an **orthogonal** matrix.

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:



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: $A = PDP^T$

$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T$

equal

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

A^{11}

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

$$\det(A-\lambda I) = \det\begin{pmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} = (3-\lambda)^2 - 1$$

$$= \lambda^2 - 6\lambda + 9 - 1$$

$$= \lambda^2 - 6\lambda + 8 = (\lambda-4)(\lambda-2)$$

$\lambda_2 = 4$

$$A - 4I = \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right)$$

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow \vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_2 \leftarrow \vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

(note: need to normalize)

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

Answer

$\lambda_1 = 2$

dot prod.

~~$\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} -b \\ a \end{bmatrix} = 0$~~

$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = 0$

$$A - 2I = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

$$\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} x_2 \qquad \vec{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{P^T}$$

Quadratic Forms

Quadratic Forms

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

Examples:

$$3x^2 + 2y^2$$

$$3x^2 + 6xy + 2y^2$$

$$3x^2 + 6xy + 2yz + 3z^2$$

~~$3x^2 + 6xy + 2y^2 + 3x$~~

Quadratic Forms and Symmetric Matrices

Fact. Every quadratic form can be represented as

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

where A is <u>symmetric</u>.

Example:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3x \\ 2y \end{bmatrix} = 3x^2 + 2y^2$$

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:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3x - 2y \\ -2x + 7y \end{bmatrix} = 3x^2 - 2xy - 2xy + 7y^2$$

$$= 3x^2 - 4xy + 7y^2$$

↑ 2×1 matrix!

Quadratic forms and Symmetric Matrices (Again)

Furthermore, we can generally say

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

Verify: $\vec{x}^T(A\vec{x}) = \sum_{i=1}^{n} x_i (A\vec{x})_i = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} A_{ij}x_j\right)$

$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i A_{ij} x_j$

$\vec{a}^T\vec{b} = \sum_{i=1}^{n} a_i b_i$

$i \rightarrow \begin{bmatrix} A_{i1} & A_{i2} & \cdots & A_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [\;]$

A Slightly more Complicated Example

$\leftarrow i=1, j=2$ entry

$$[x\ y\ z]\, A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ -1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

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

$$x^2 + 3y^2 + 5z^2 + (2+2)xy + (-1-1)xz$$

$$x^2 + 3y^2 + 5z^2 + 4xy - 2xz$$

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

$$A = \begin{bmatrix} 5 & -0.5 & 0 \\ -0.5 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

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 $A_{ij} = A_{ji} = \dfrac{\alpha}{2}$

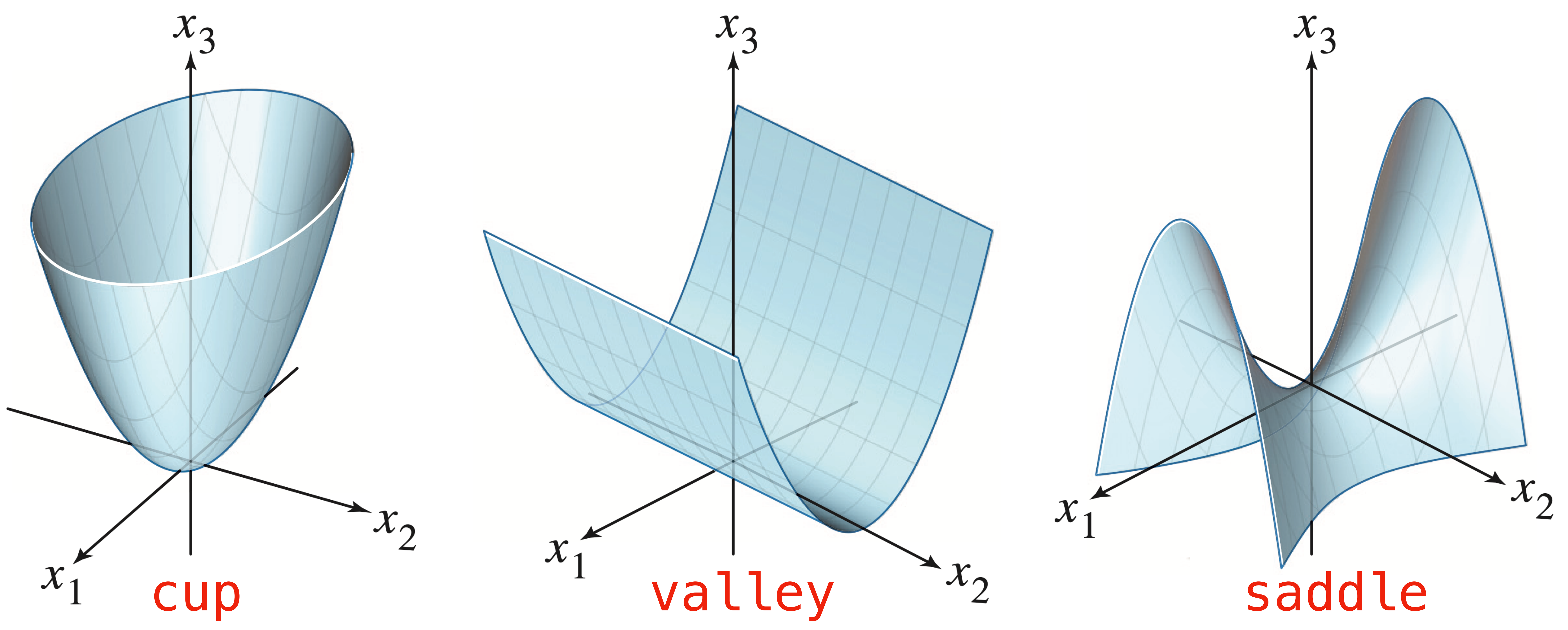
Practice Problem

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

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

$$A = \begin{bmatrix} 1 & 0 & 3.5 & 0 \\ 0 & 3 & 0 & 0 \\ 3.5 & 0 & 0 & -1 \\ 0 & 0 & -1 & -6 \end{bmatrix}$$

Shapes of of Quadratic Forms

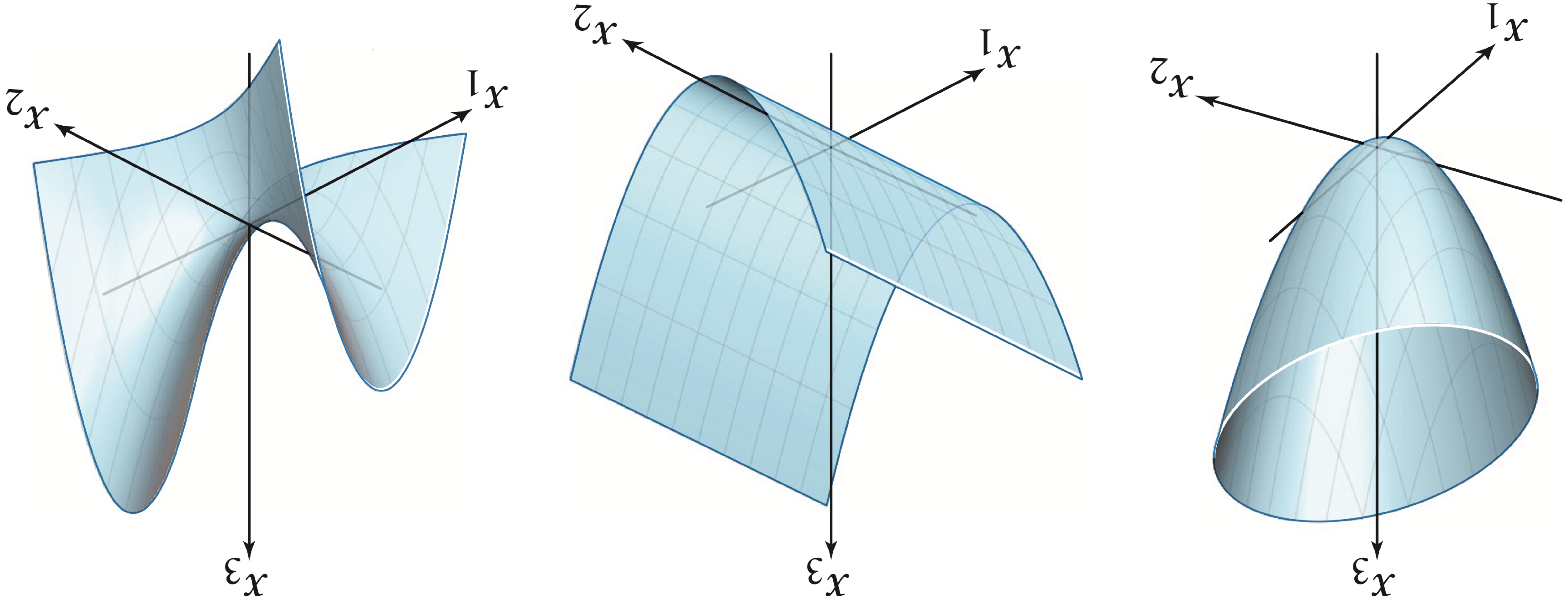


There are essentially three possible shapes (six if you include the negations).

Can we determine what shape it will be mathematically?

Linear Algebra and its Applications, Lay, Lay, McDonald

Shapes of of Quadratic Forms

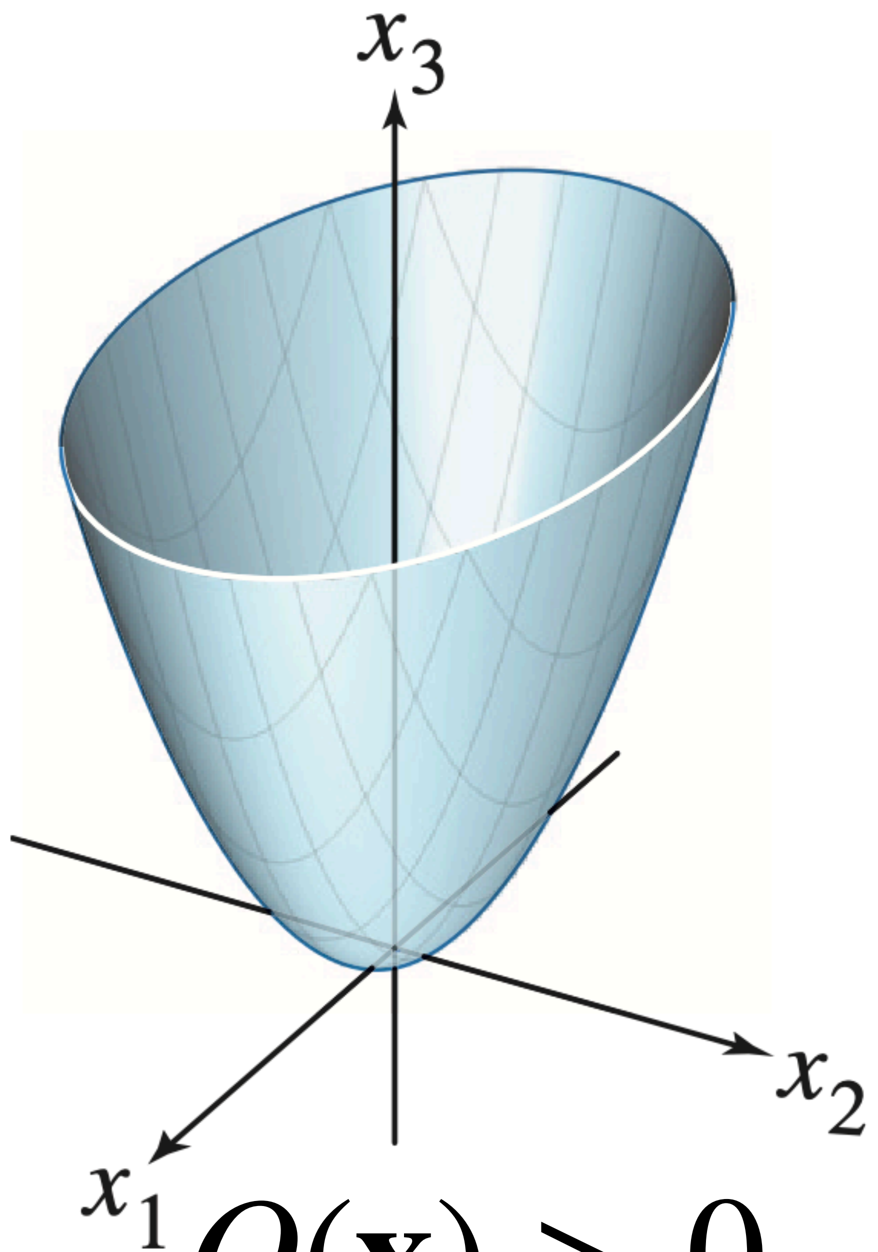


There are essentially three possible shapes (six if you include the negations).

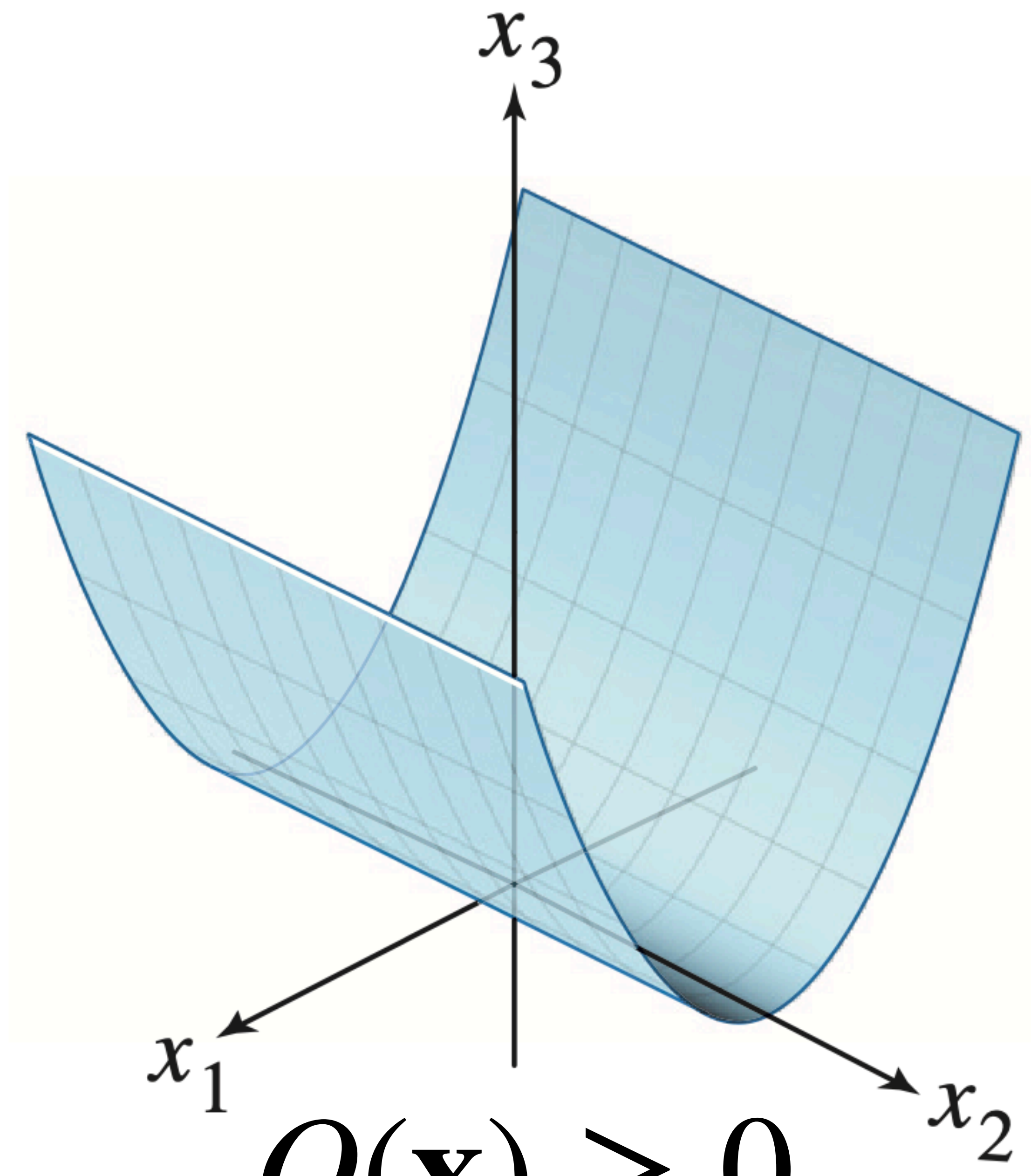
Can we determine what shape it will be mathematically?

Linear Algebra and its Applications, Lay, Lay, McDonald

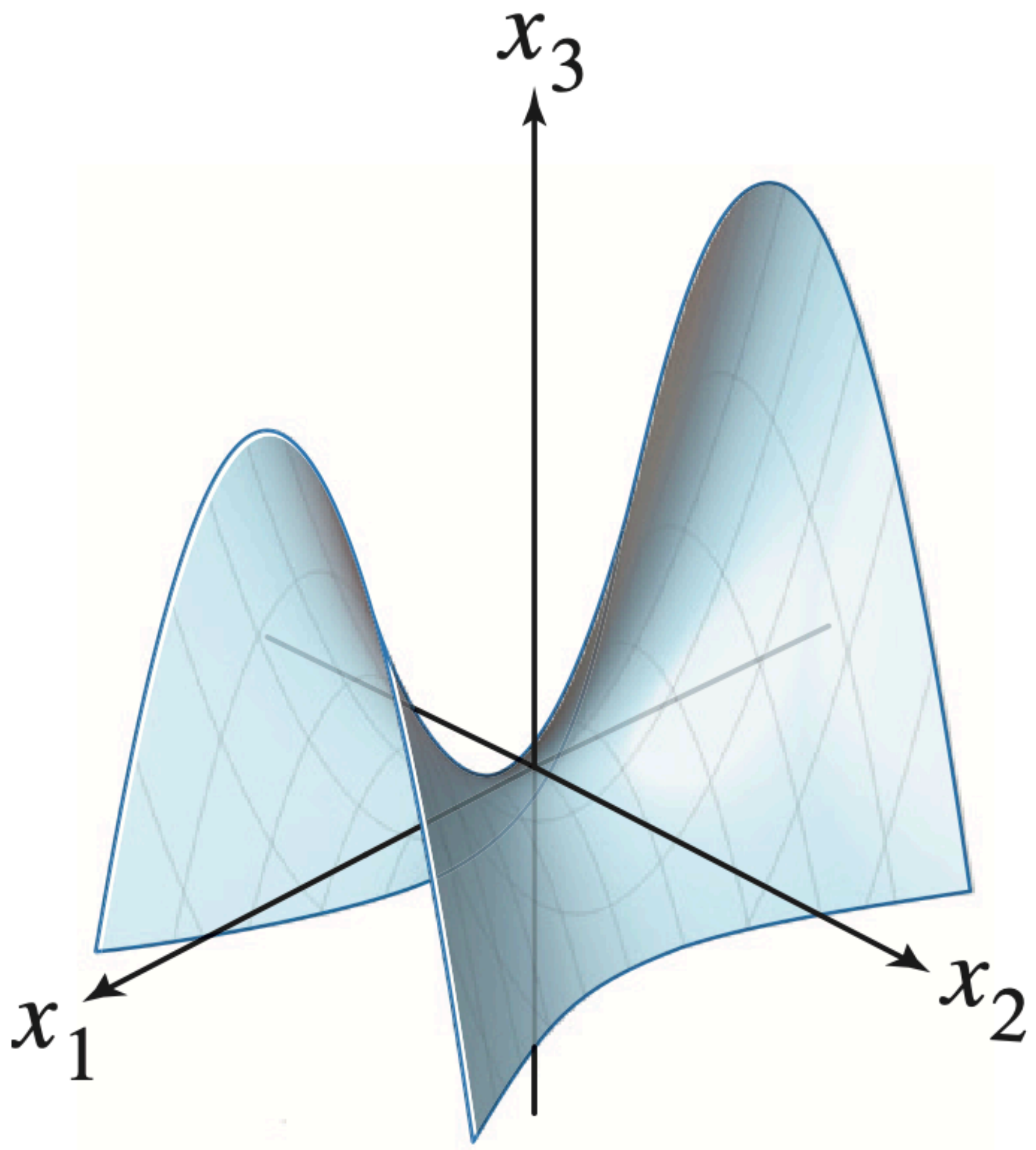
Definiteness



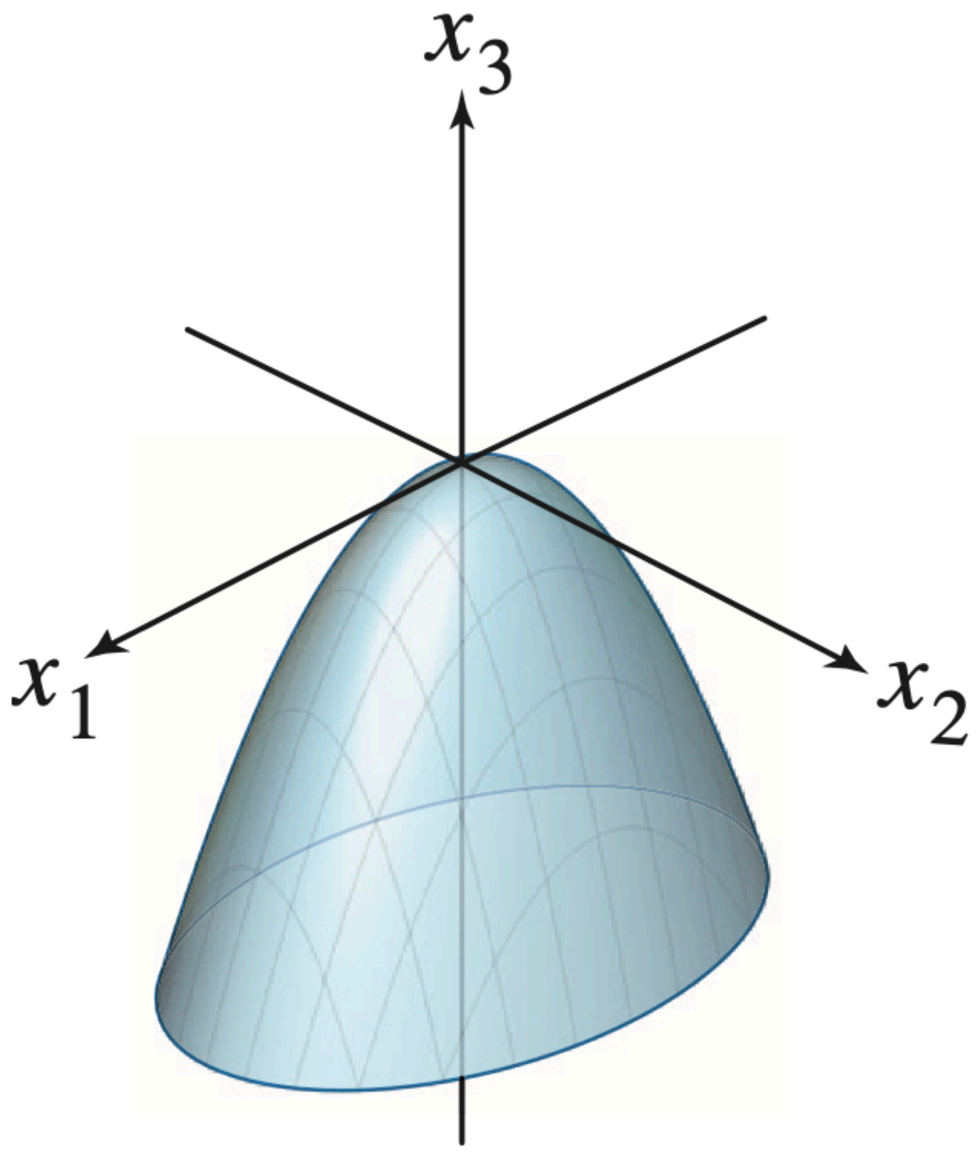
$Q(\mathbf{x}) > 0$



$Q(\mathbf{x}) \geq 0$



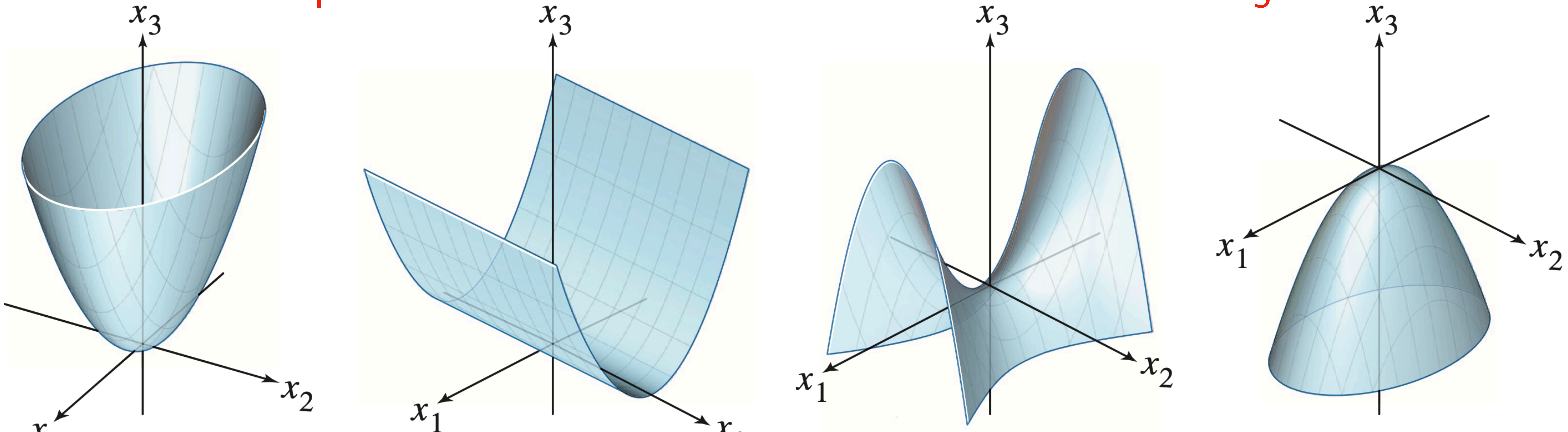
$Q(\mathbf{x})$ can be + & –



$Q(\mathbf{x}) < 0$

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

Definiteness



positive semidefinite — negative definite

$Q(\mathbf{x}) > 0$ — $Q(\mathbf{x}) \geq 0$ — $Q(\mathbf{x})$ can be + & – — $Q(\mathbf{x}) < 0$

positive definite — indefinite

For $\mathbf{x} \neq \mathbf{0}$, 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** ≡ all positive eigenvalues
- **positive semidefinite** ≡ all nonnegative eigenvalues
- **indefinite** ≡ positive and negative eigenvalues
- **negative definite** ≡ all negative eigenvalues

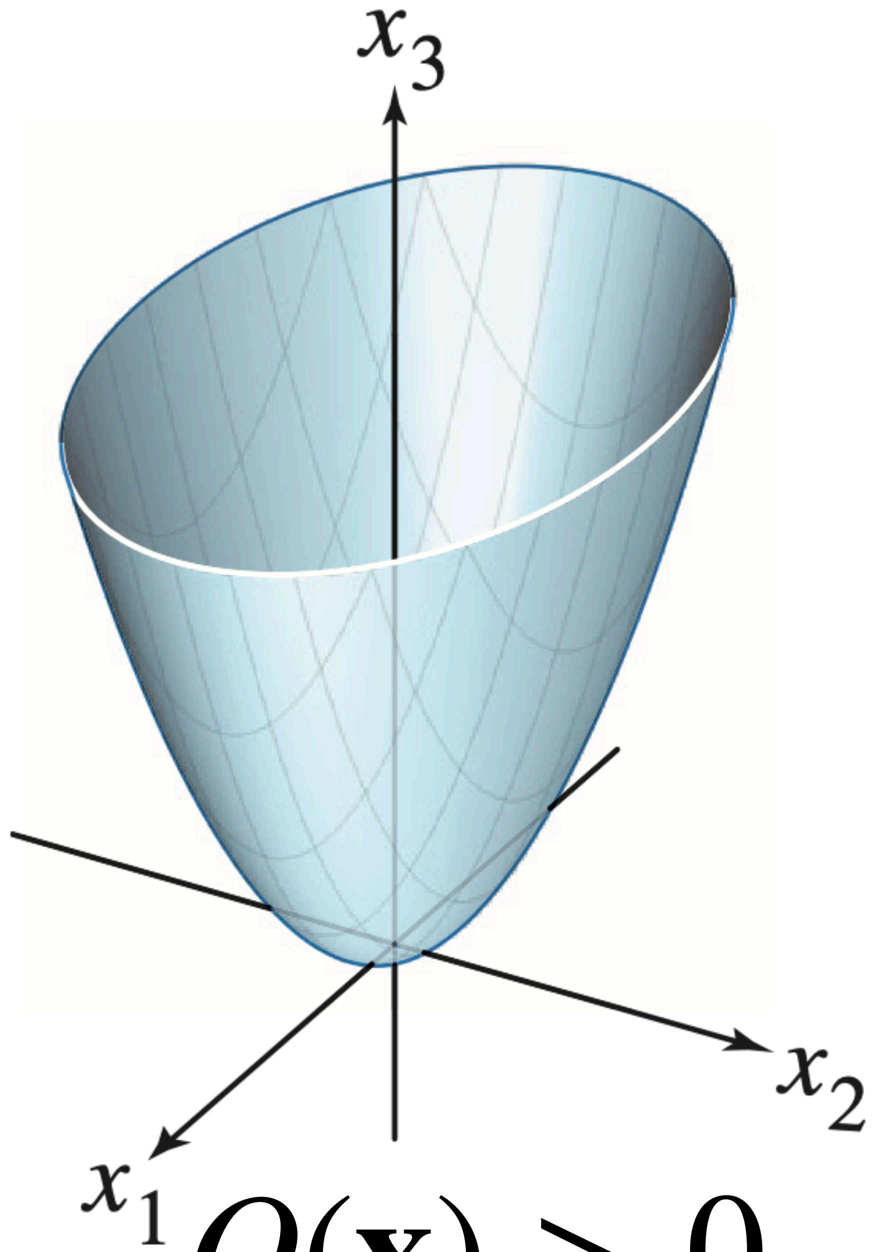
Definiteness

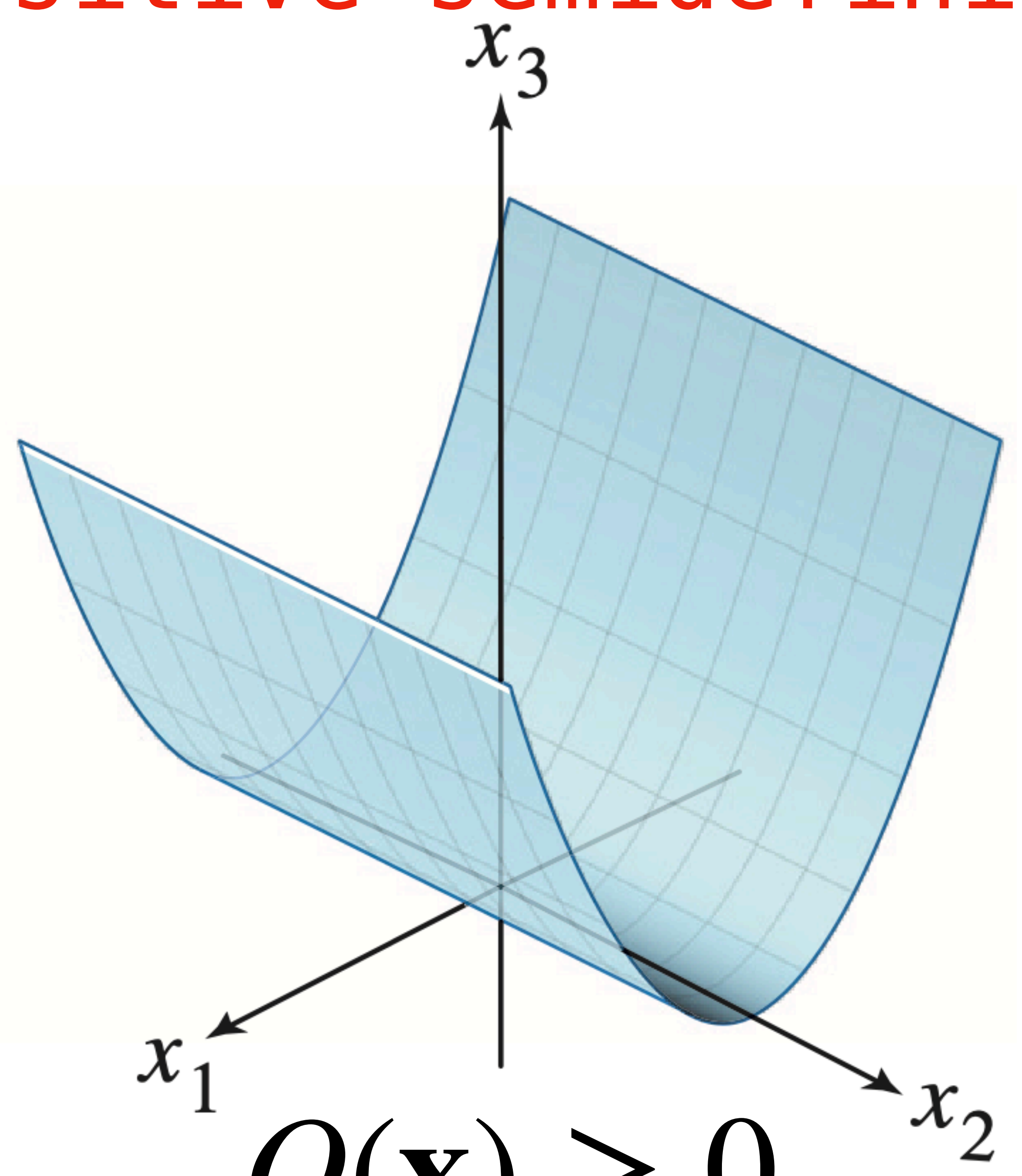
all nonneg. eigenvals

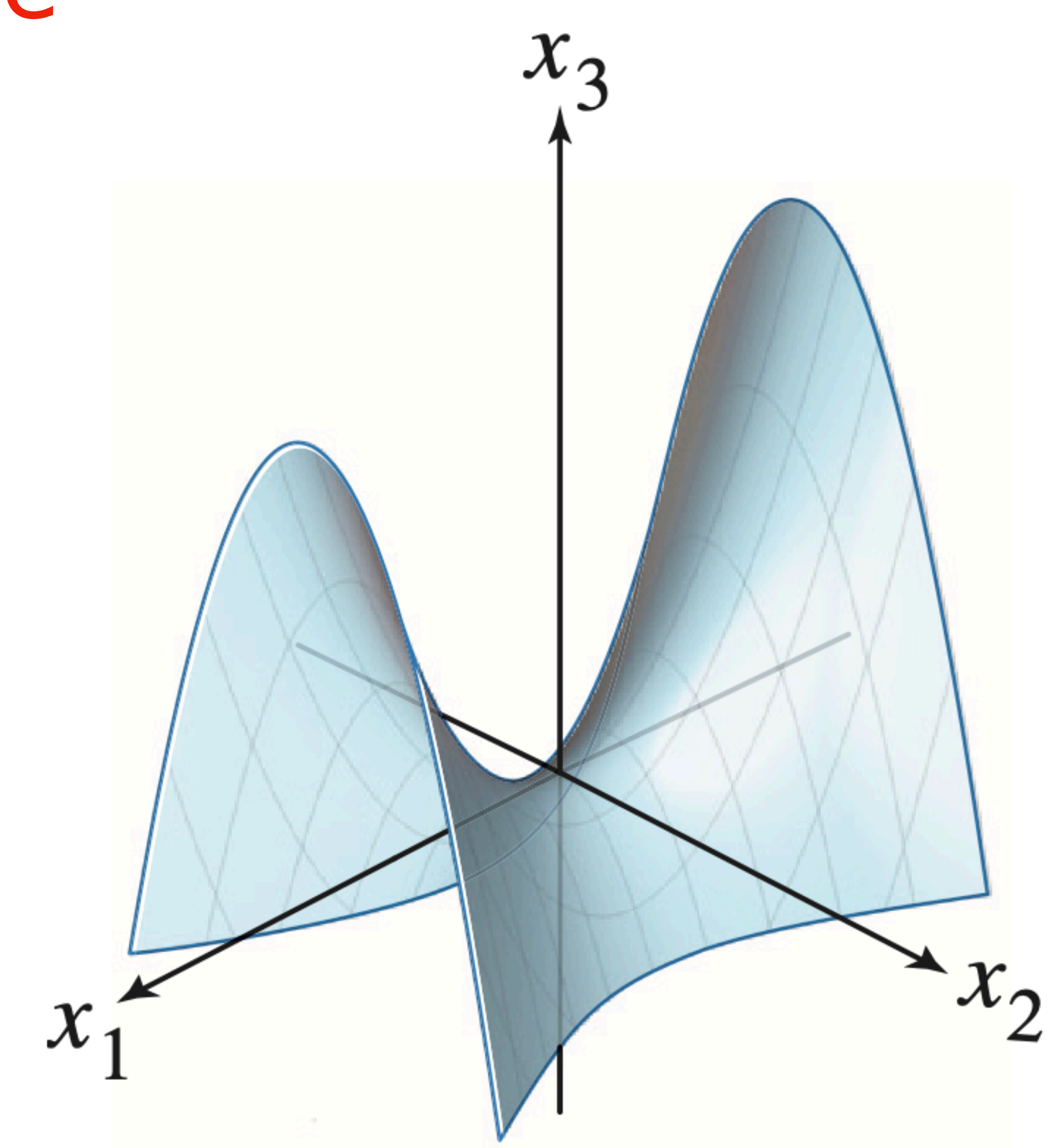
positive semidefinite

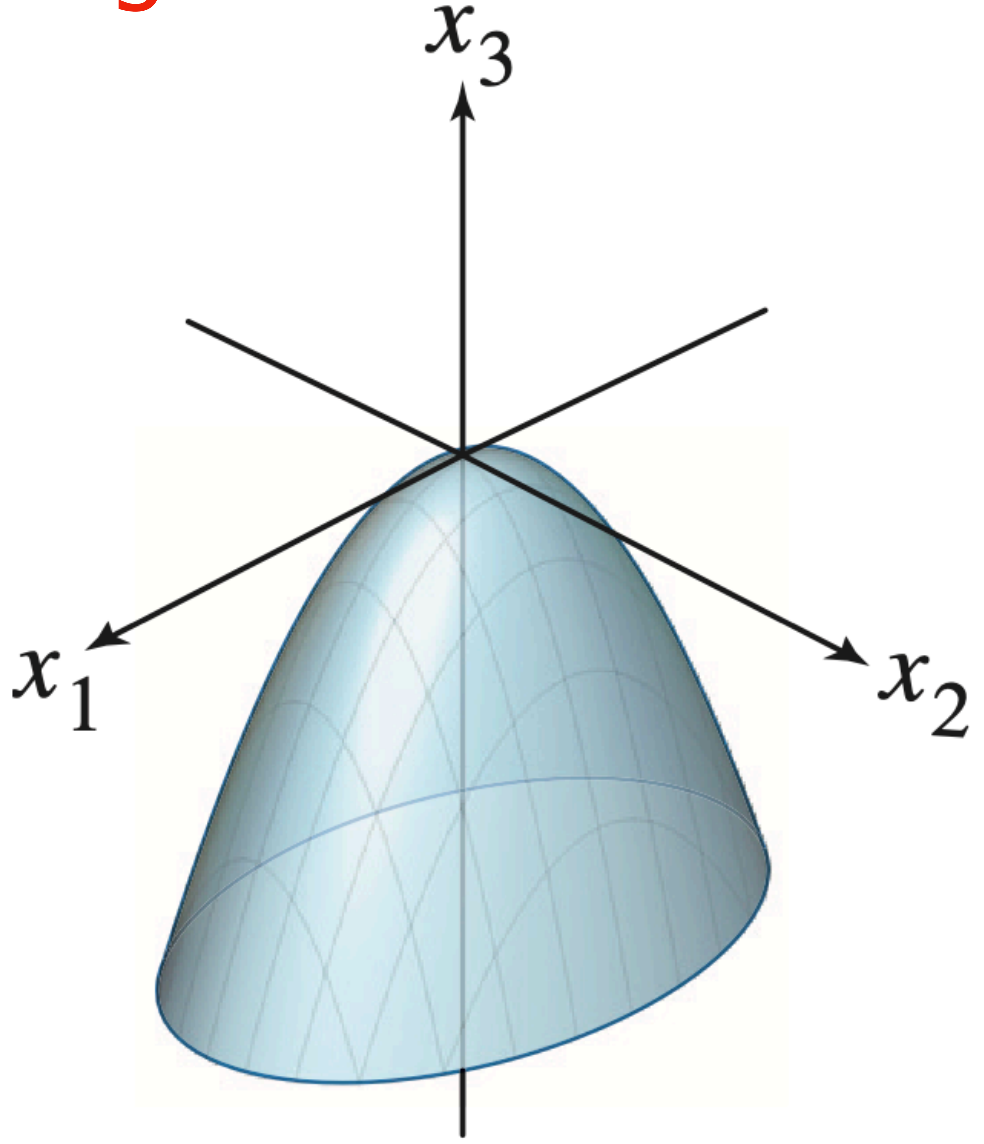
all neg. eigenvals

negative definite









$Q(\mathbf{x}) > 0$

positive definite

all pos. eigenvals

$Q(\mathbf{x}) \geq 0$

$Q(\mathbf{x})$ can be + & −

indefinite

pos. and neg. eigenvals

$Q(\mathbf{x}) < 0$

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

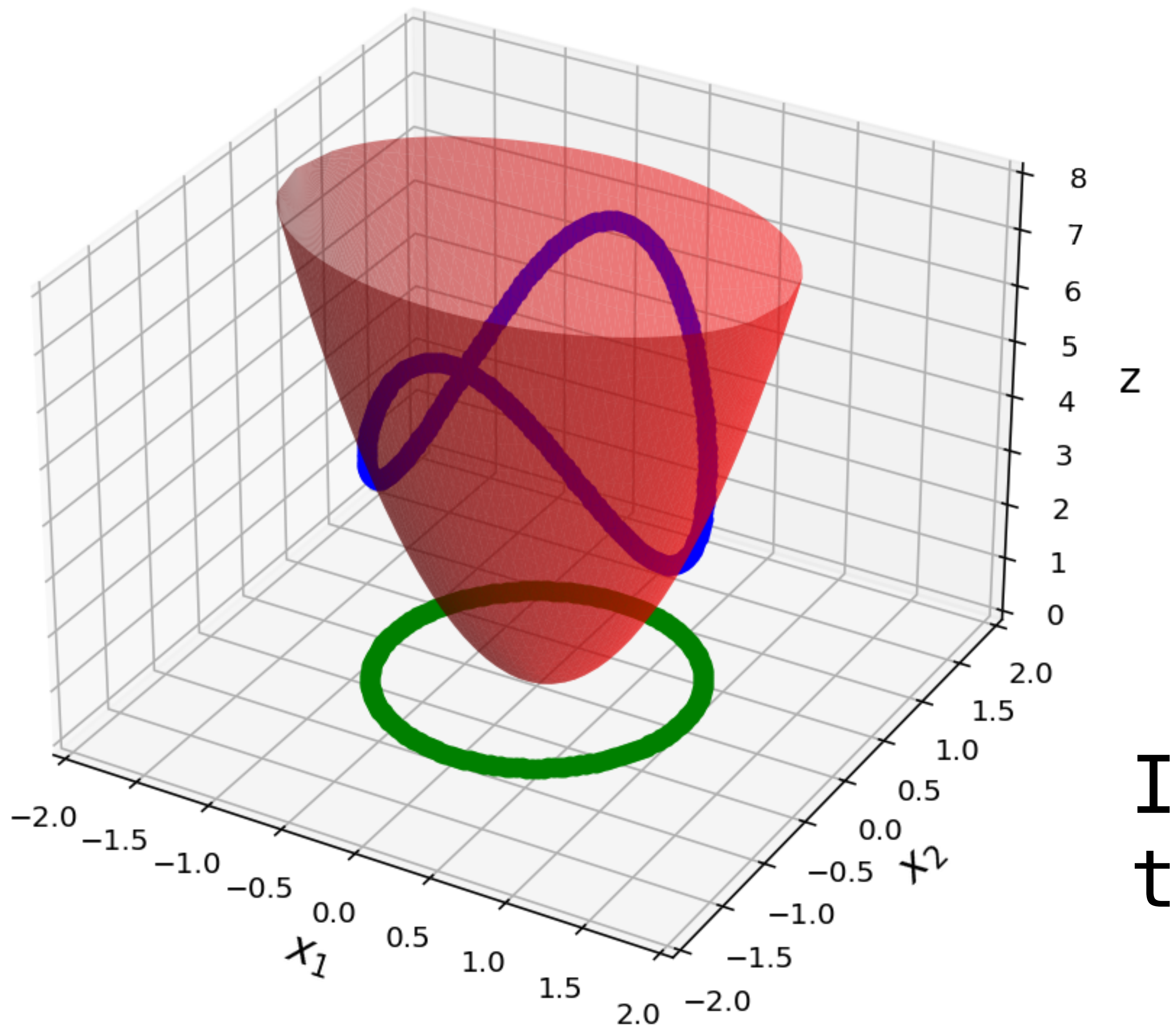
$$\min_{\mathbf{v} \in X} f(\mathbf{v})$$

(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

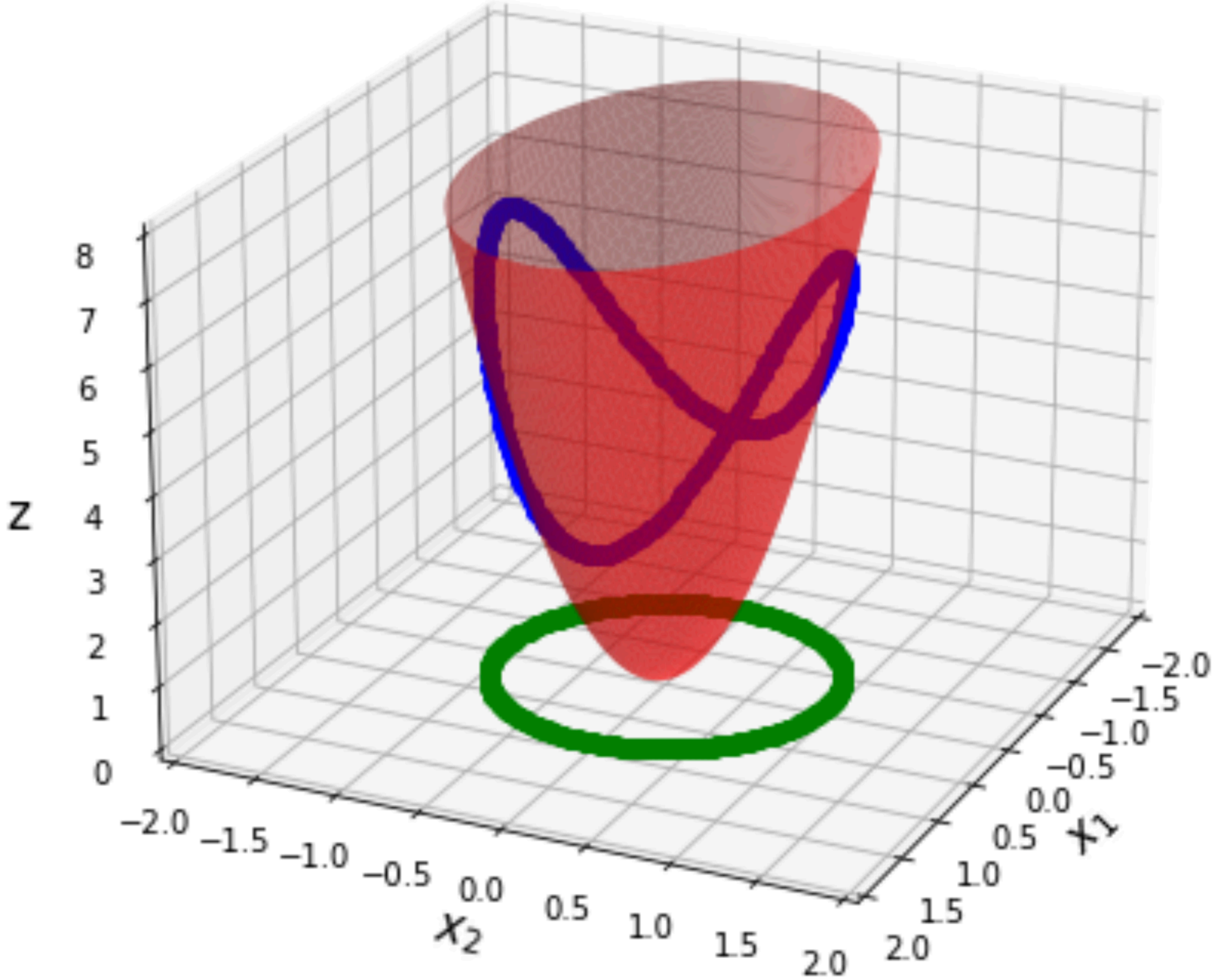
mini/maximize $\mathbf{x}^T A \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$



It's common to constraint to unit vectors.

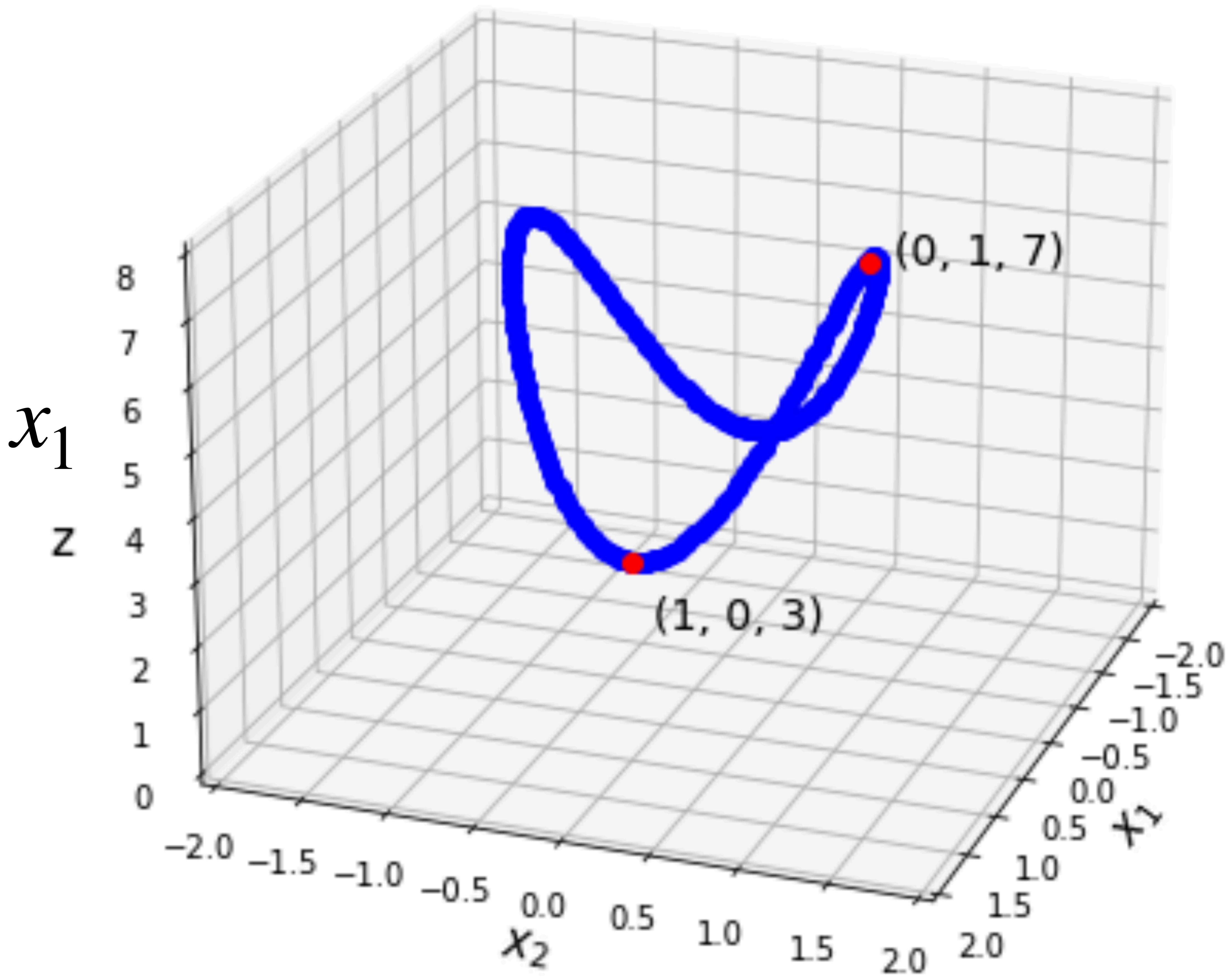
Example: $3x_1^2 + 7x_2^2$

What are the min/max values?:



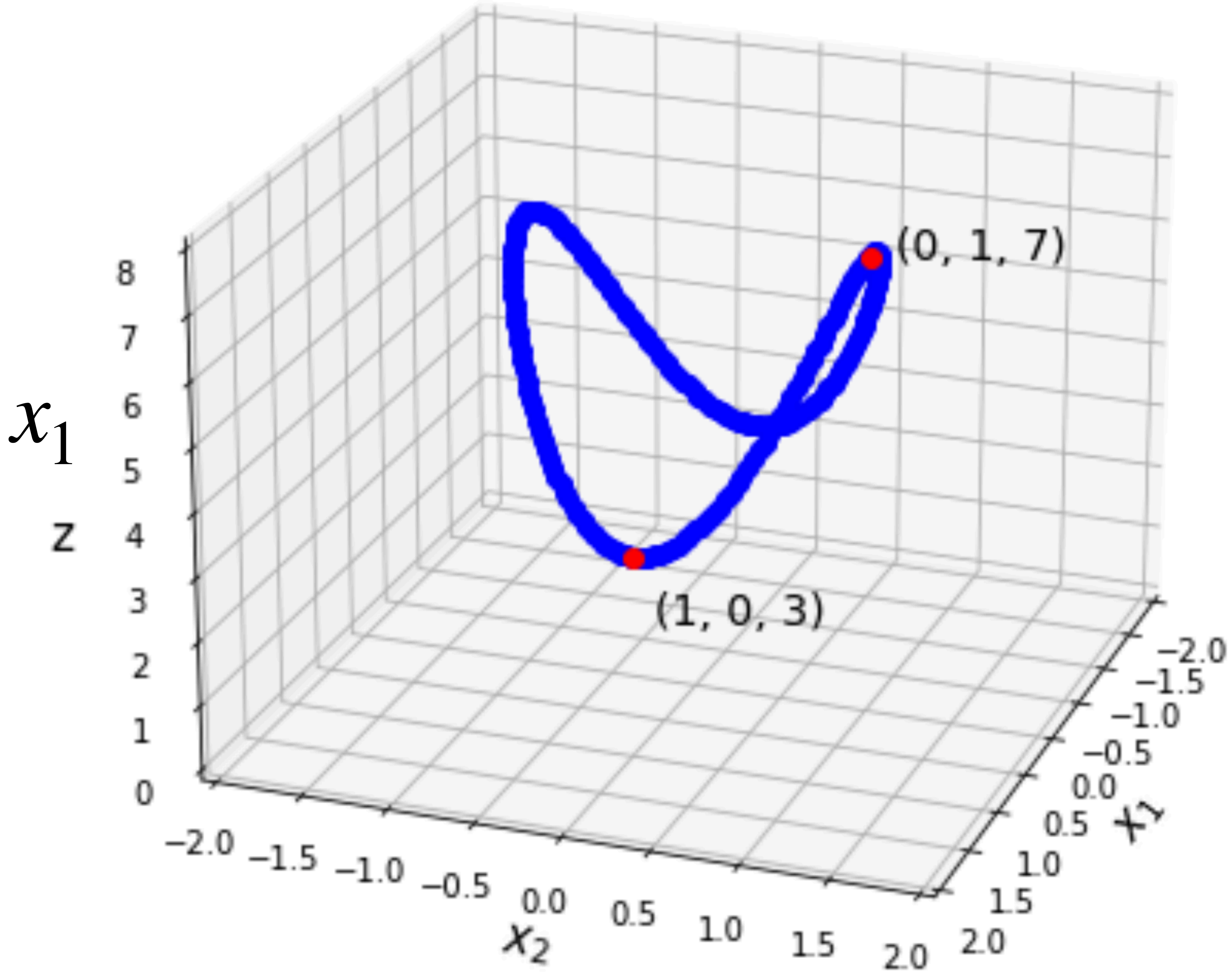
Example: $3x_1^2 + 7x_2^2$

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

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

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

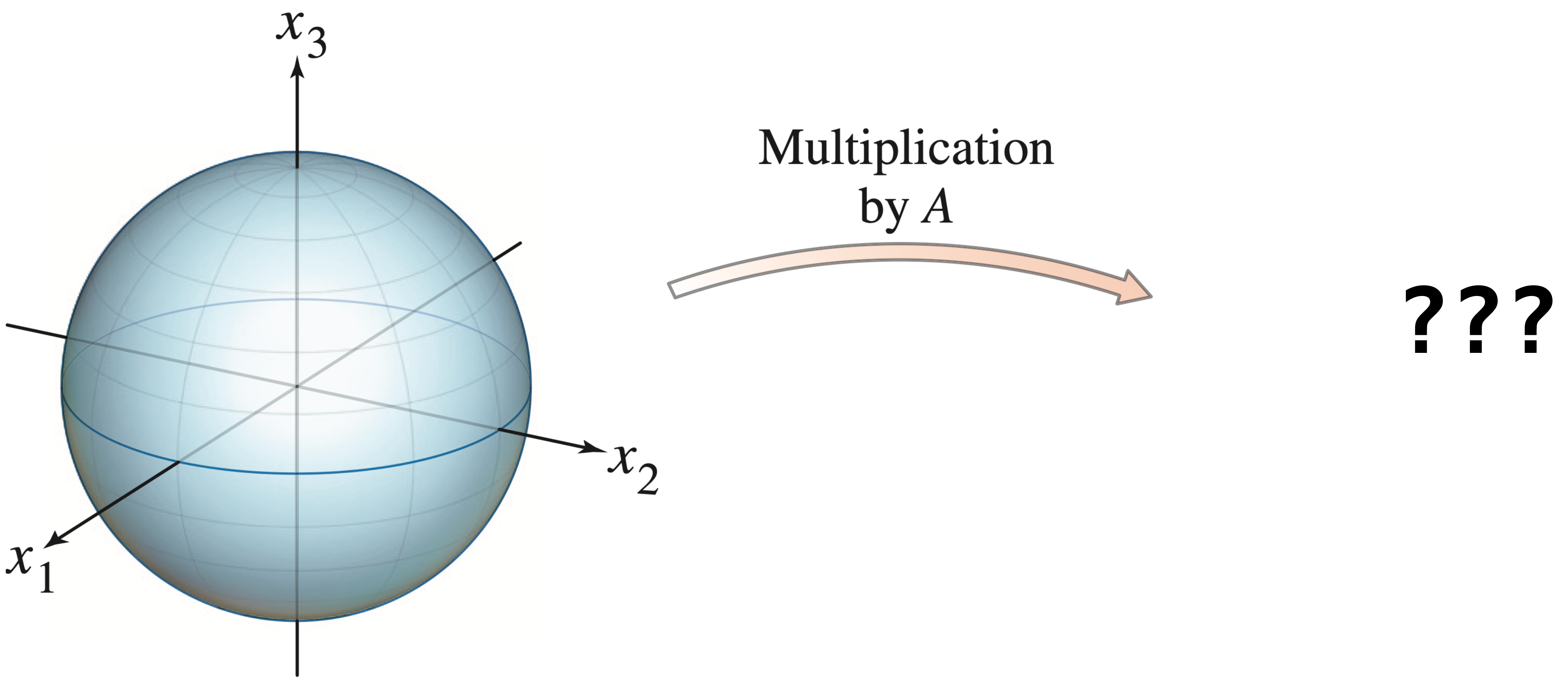
$$Q(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 4x_2x_3 + x_3^2$$

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

Singular Value Decomposition (Looking Ahead)

Question

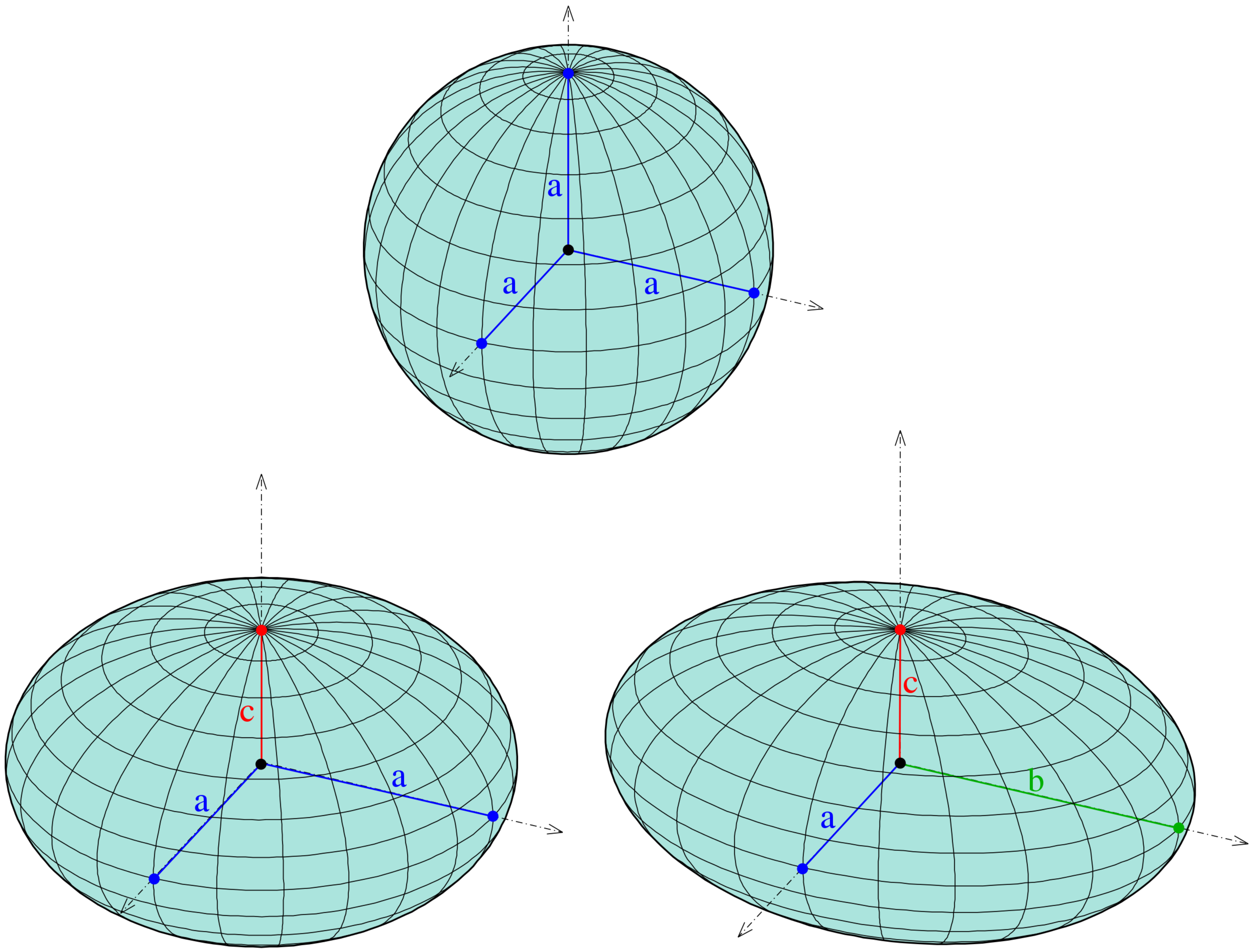
What shape is a the unit sphere after a linear transformation?



Ellipsoids

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

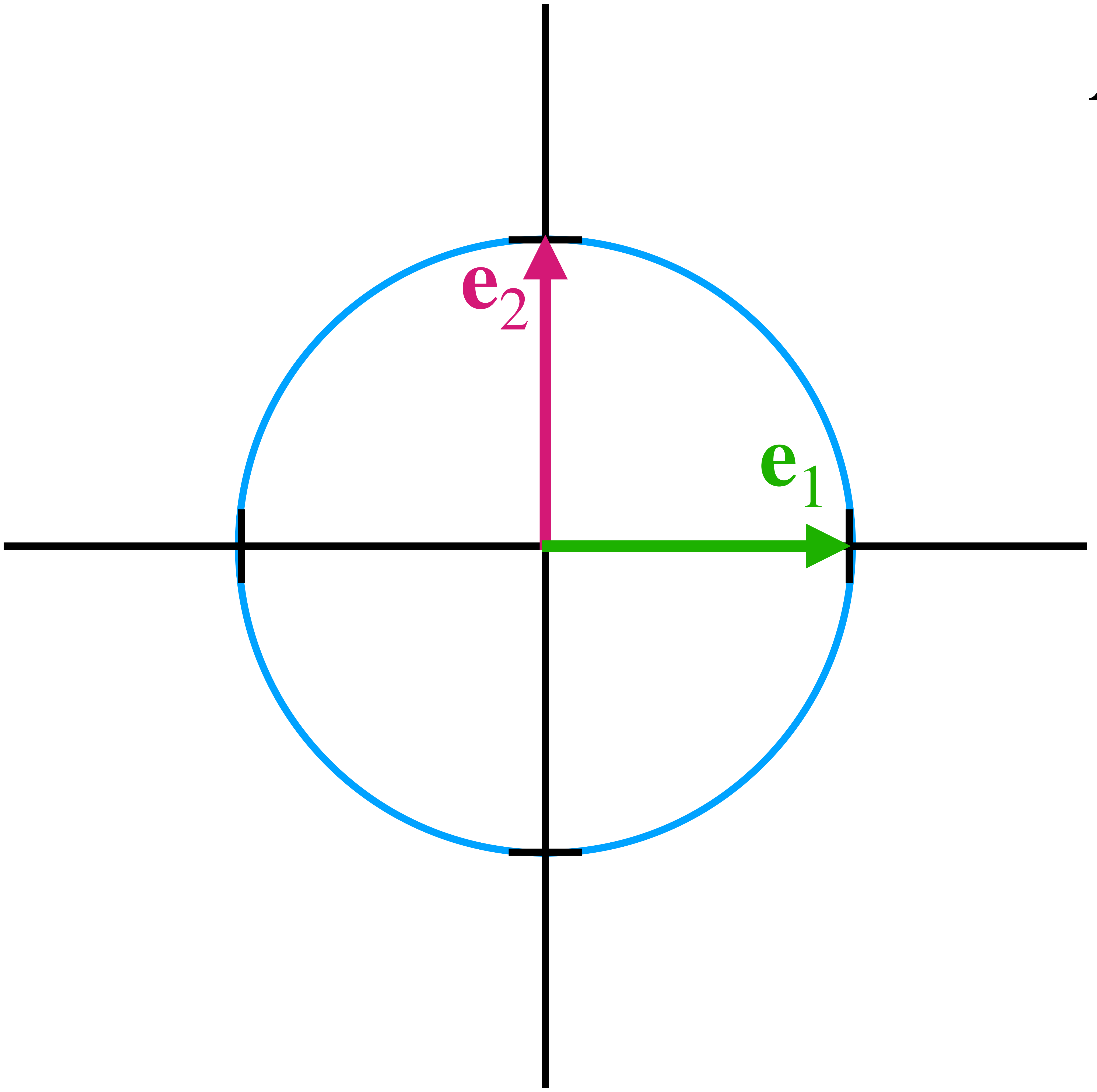
Linear transformations maps <u>spheres</u> to <u>ellipsoids</u>.

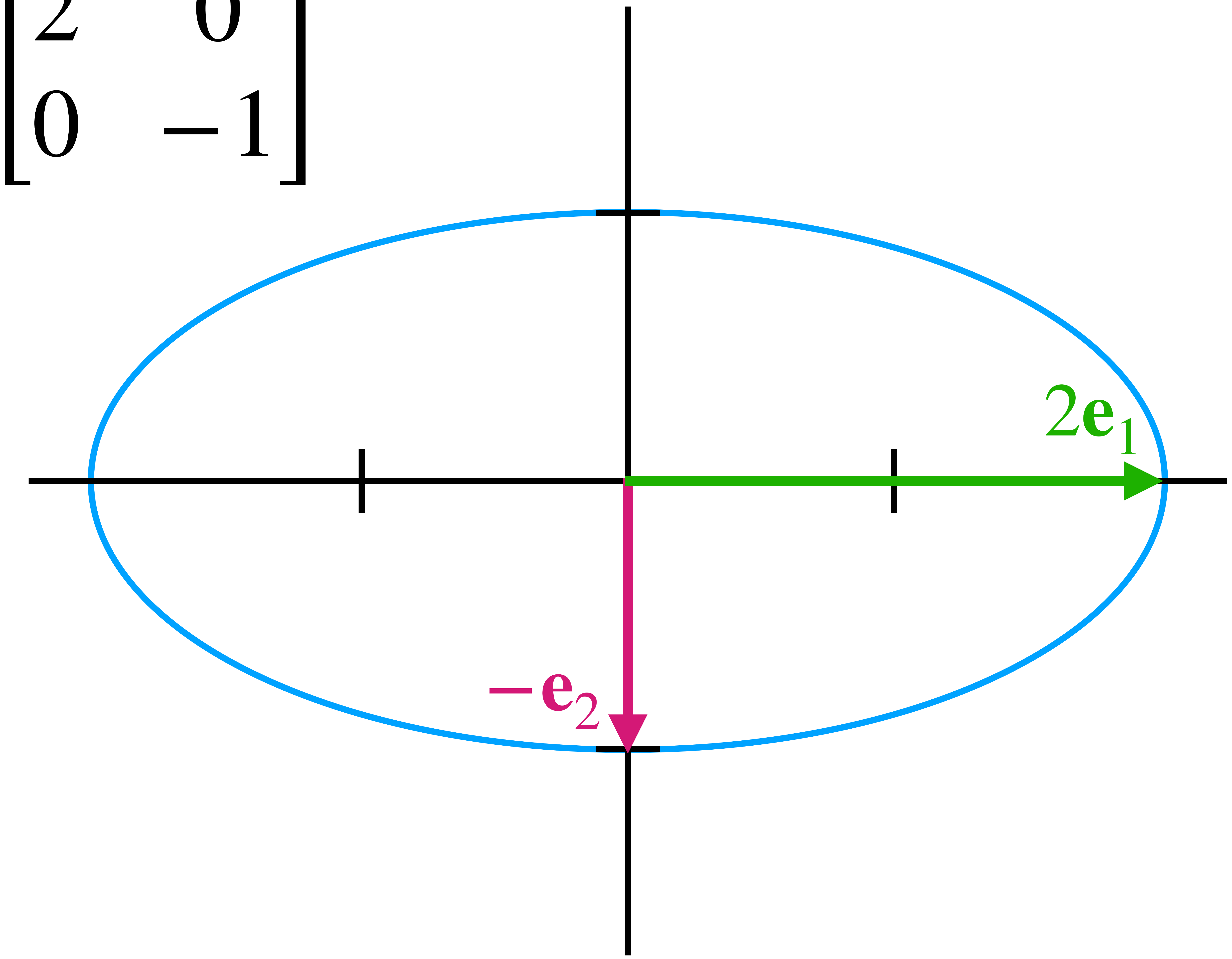


https://commons.wikimedia.org/wiki/File:Ellipsoide.svg

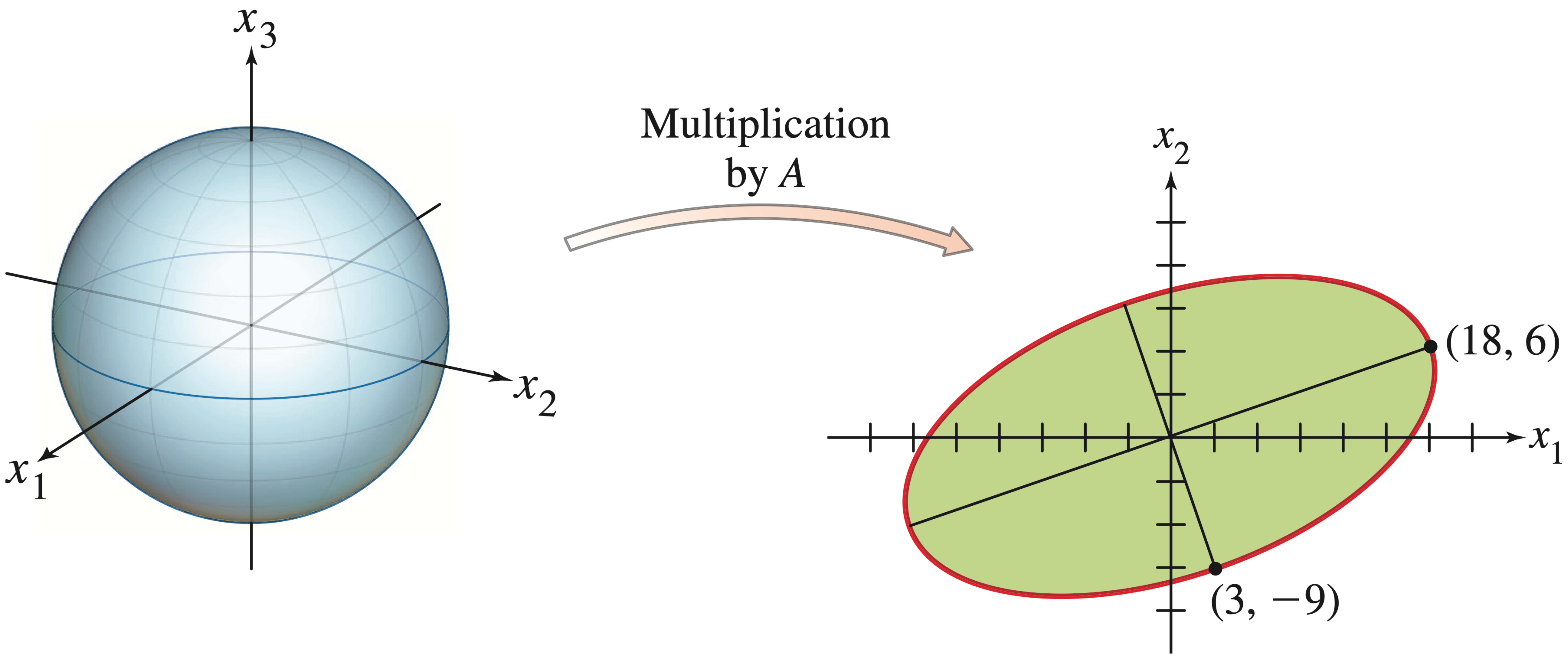
Simple Example : Scaling Matrices

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

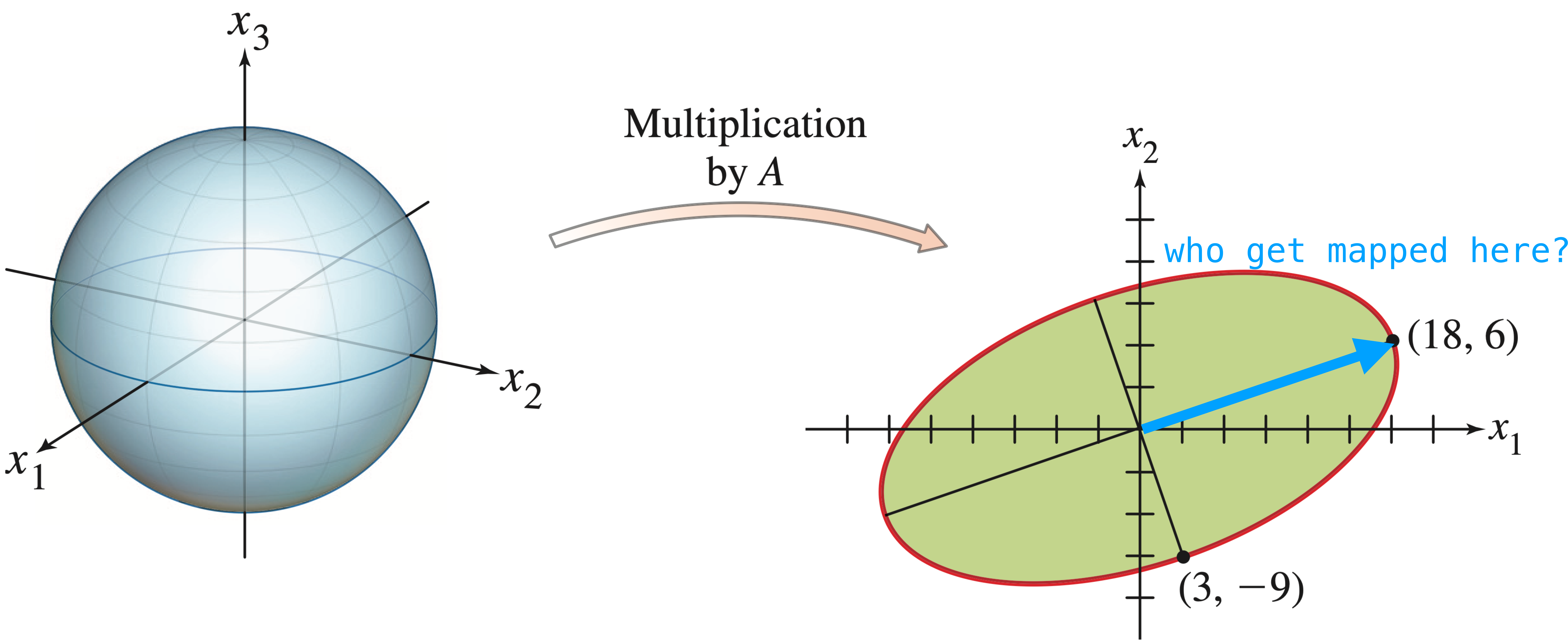




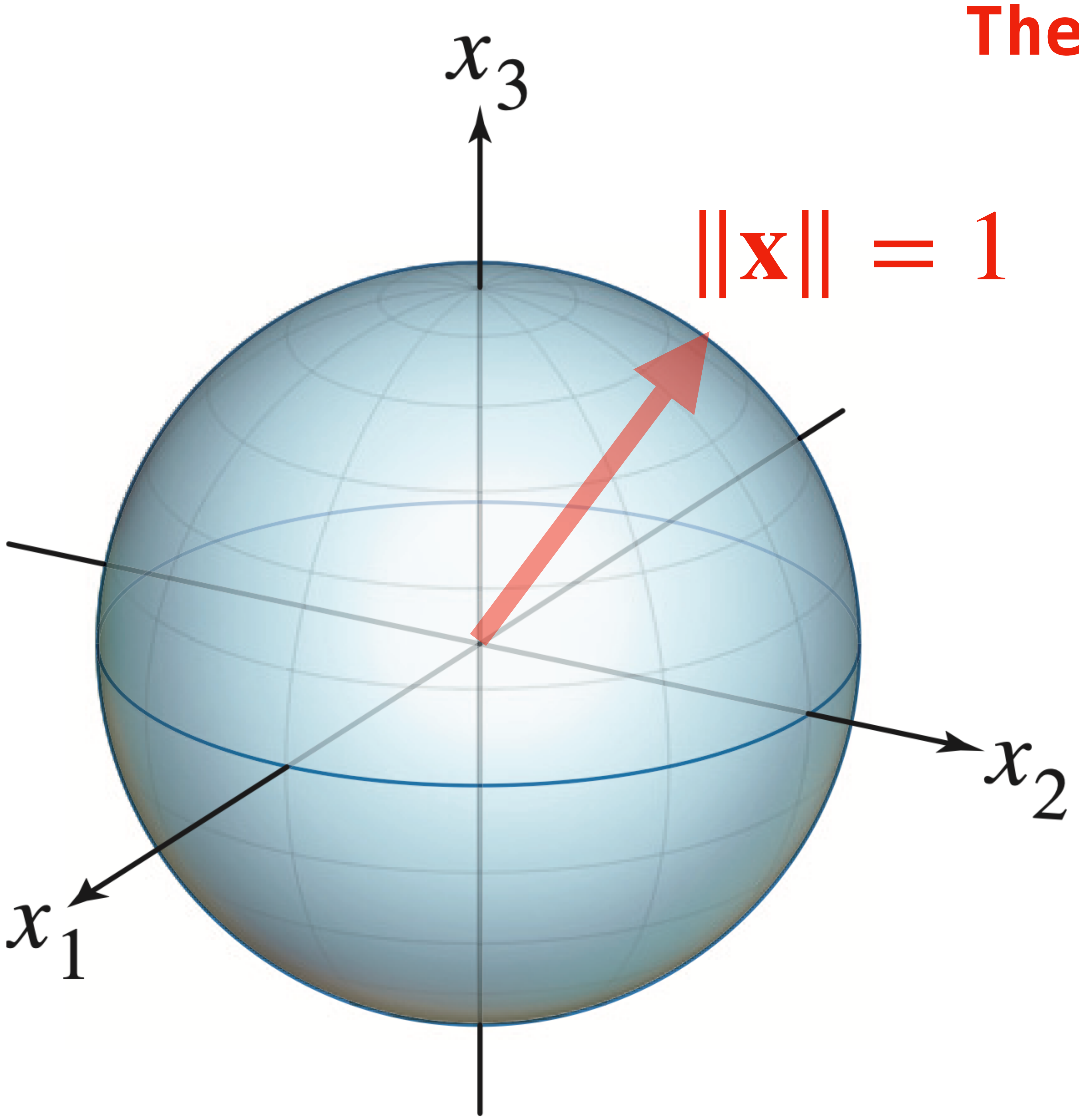
The Picture



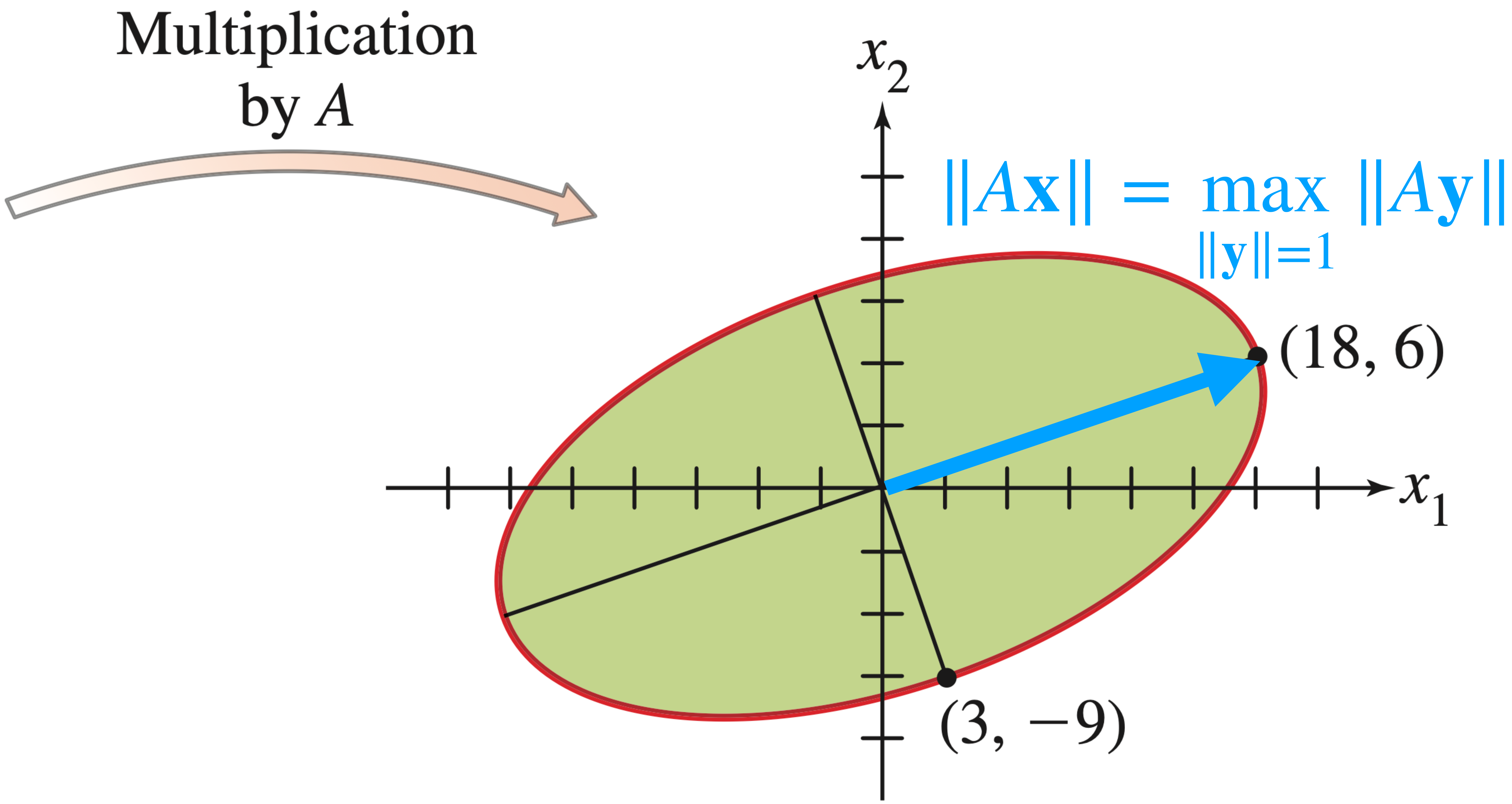
The Picture



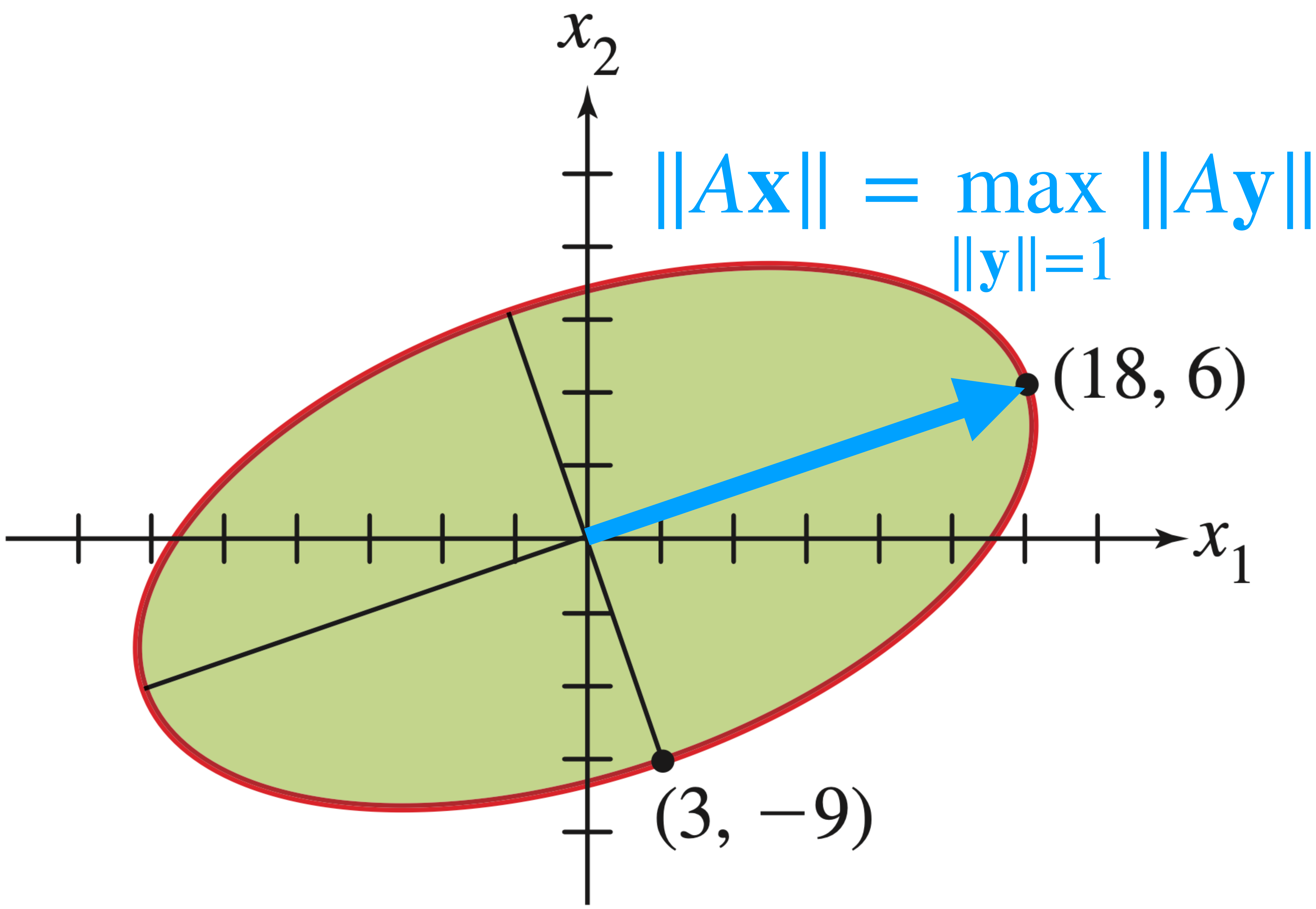
The Picture



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

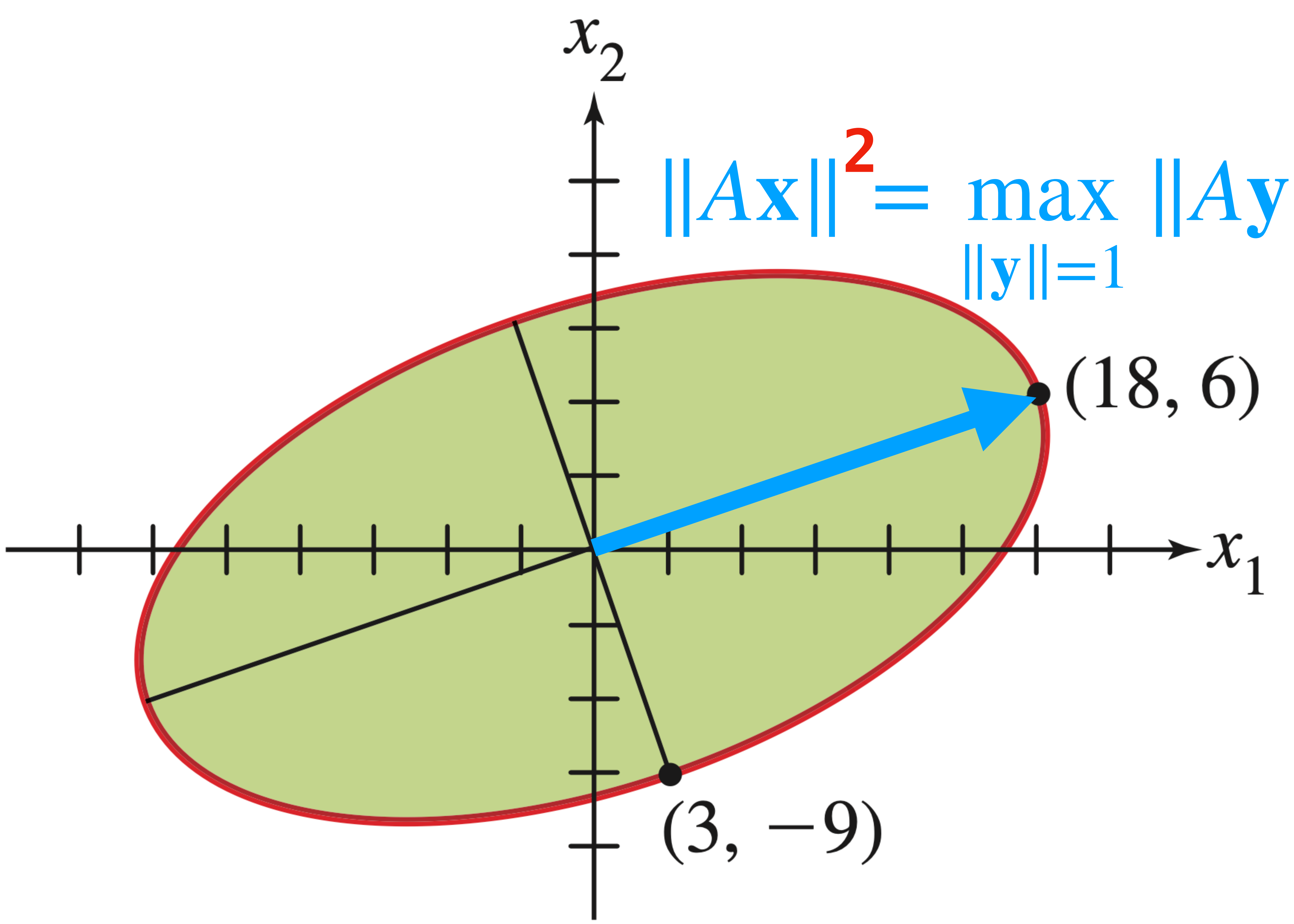


The Picture



This is not a quadratic form...

The Picture



$$\|A\mathbf{x}\|^2 = \max_{\|\mathbf{y}\|=1} \|A\mathbf{y}\|^2$$

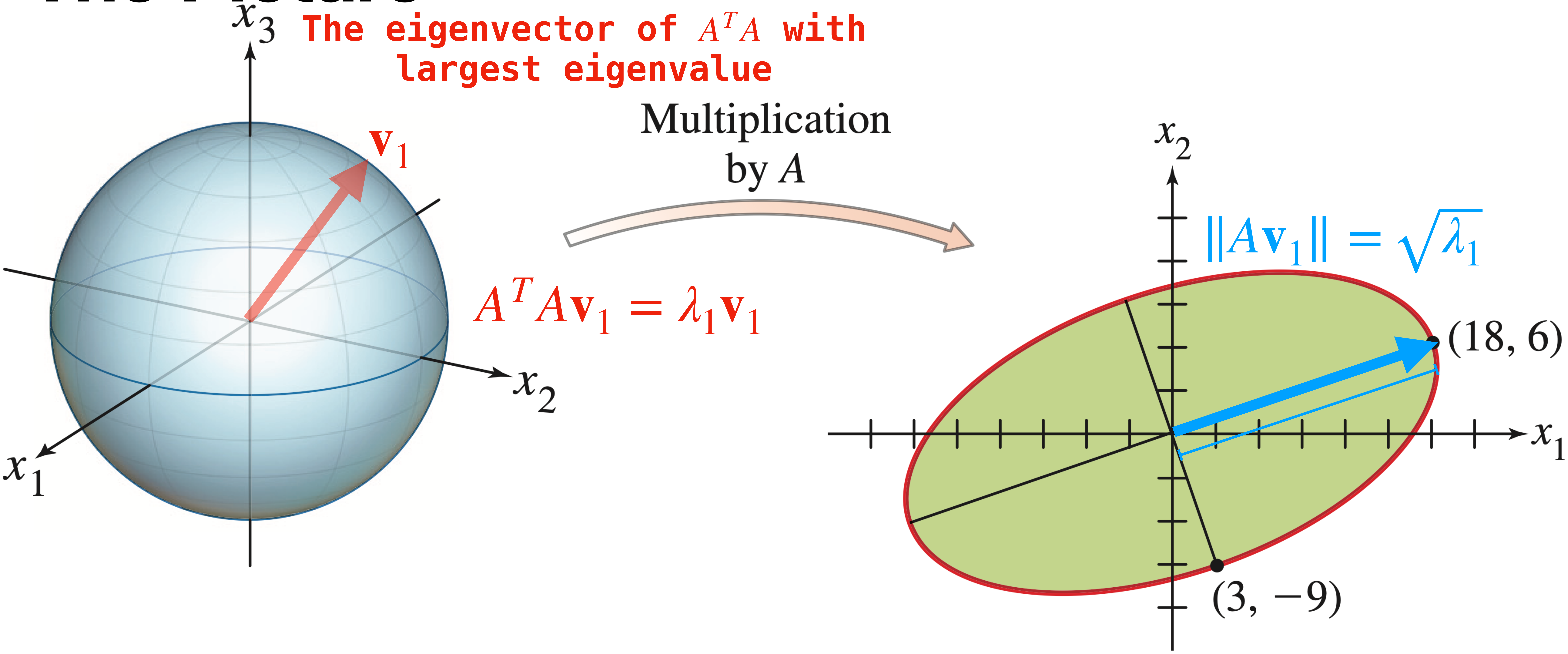
But this is.

This is not a quadratic form...

A Quadratic Form

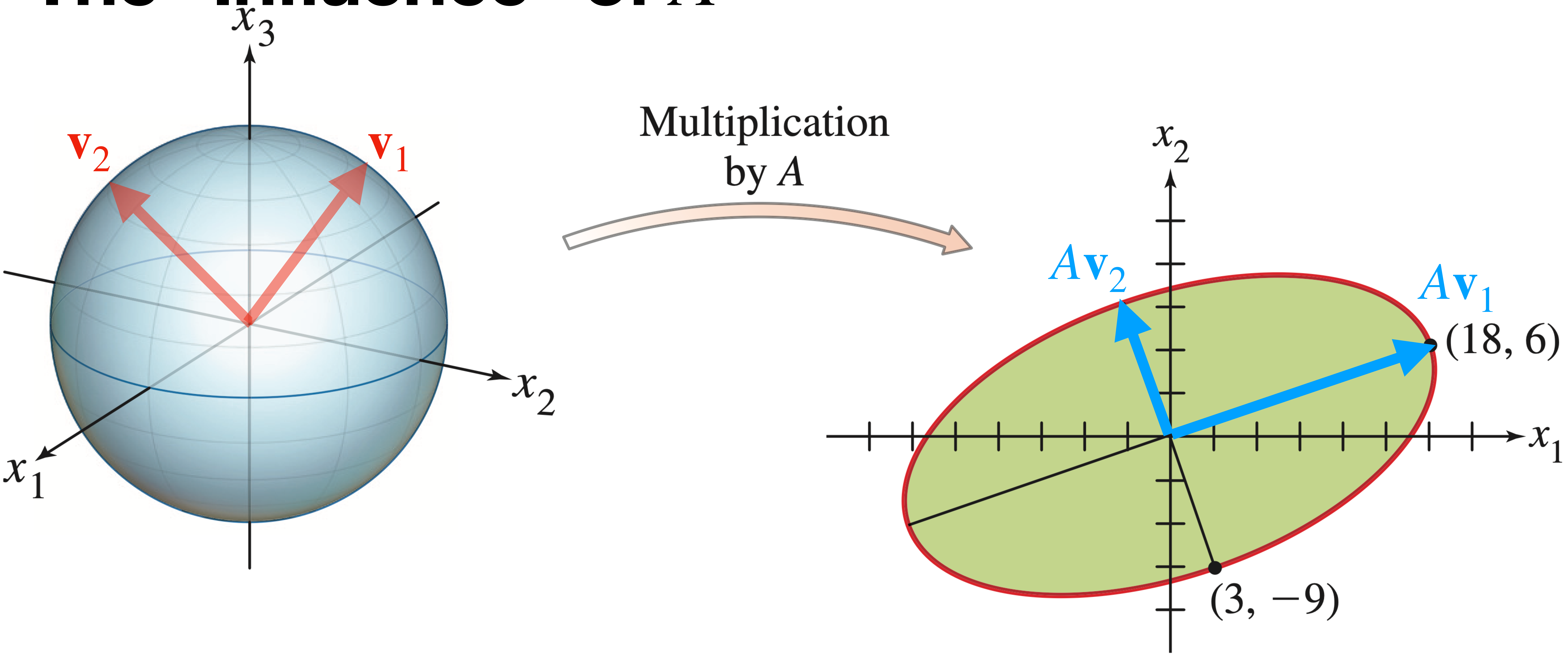
What does $\|A\mathbf{x}\|^2$ look like?:

The Picture



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

The "Influence" of A



$\mathbf{v}_1$ **is "most affected" by A and $\mathbf{v}_2$ is "least affected"**

Properties of $A^T A$

Properties of $A^T A$

» It's symmetric.

Properties of A^TA

- 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.**
- It's eigenvalues are nonnegative.

Properties of $A^T A$

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

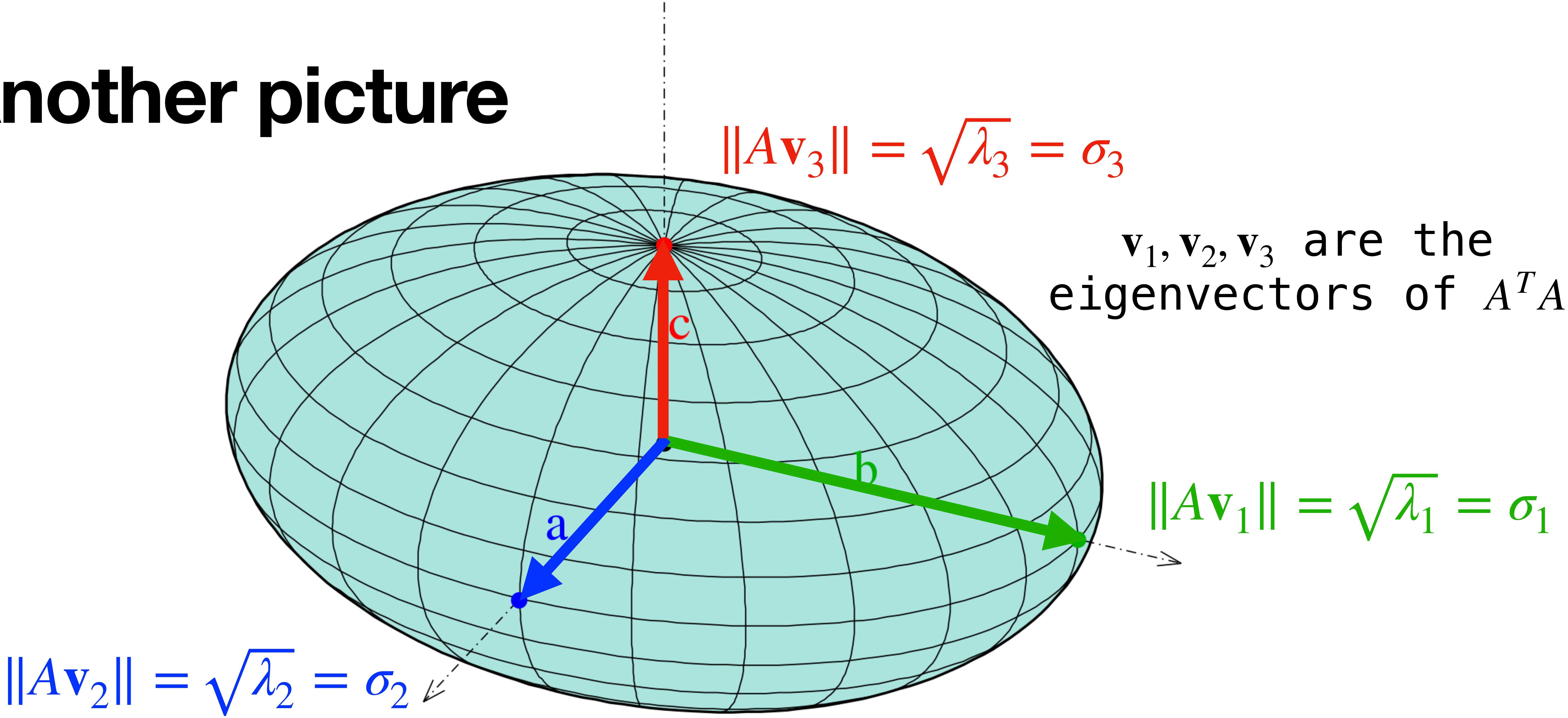
Singular Values

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

$$\sigma_1 \geq \sigma_2 \ldots \geq \sigma_n \geq 0$$

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

Another picture



$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are the eigenvectors of A^TA

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

https://commons.wikimedia.org/wiki/File:Ellipsoide.svg

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

So <u>every</u> $m \times n$ matrix has n singular values.