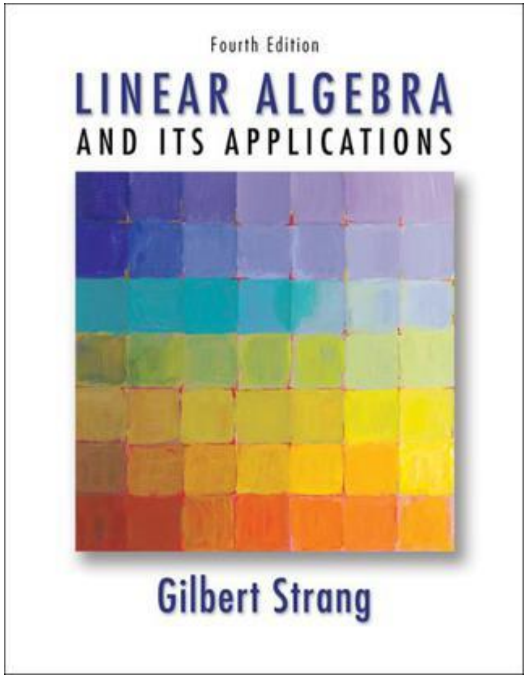
Linear Algebra



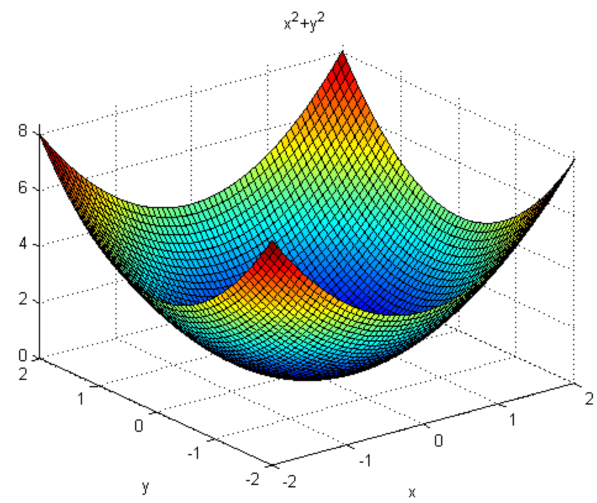
Instructor: Jing YAO

6 Positive Definite Matrices (正定矩阵)

6.2

TESTS FOR POSITIVE DEFINITENESS (正定性的判定)

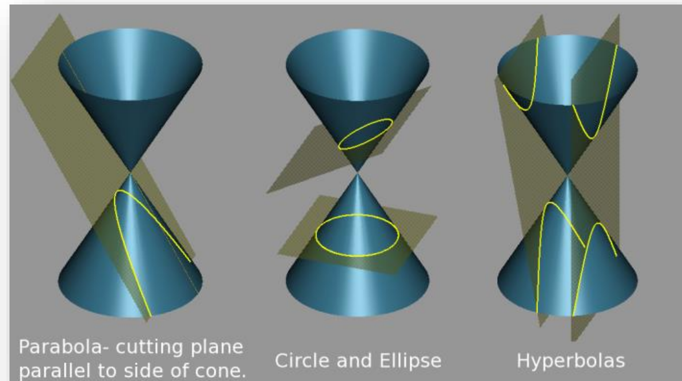
- Tests for Positive Definiteness
- Positive Semidefinite Matrices
- The Principal Axes Theorem
- The Law of Inertia



A quadratic form (二次型) is a homogeneous polynomial of degree two in a number of variables (二次齐次多项式).

For example, $4x^2 + 2xy - 3y^2$ is a quadratic form in the variables x and y.

quadratic curve
(二次曲线)

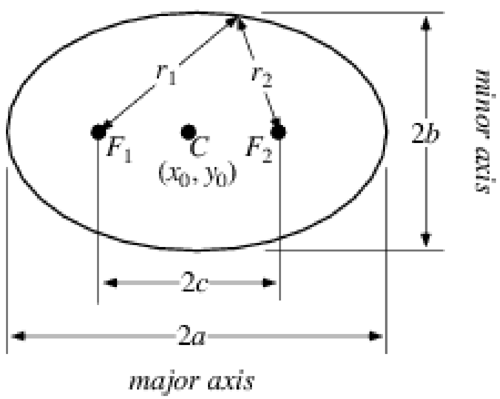


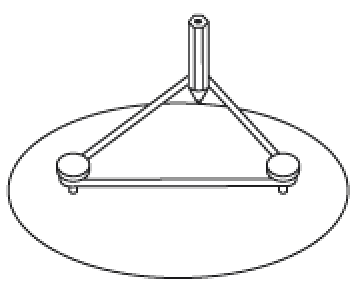
$$\boxed{\frac{x^2}{a^2}+\frac{y^2}{b^2}}=1$$

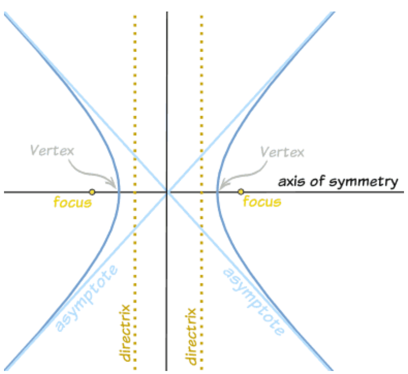
Ellipse
(椭圆)

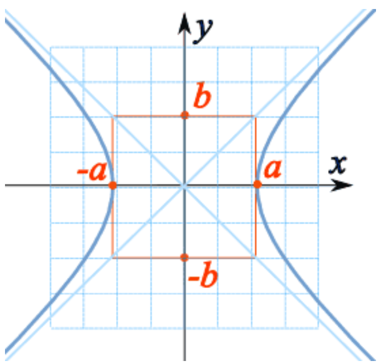
$$\boxed{\frac{x^2}{a^2}-\frac{y^2}{b^2}}=1$$

Hyperbola
(双曲线)









I. Tests for Positive Definiteness (正定性的判定)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, and the corresponding quadratic form is

$f(x, y) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = ax^2 + 2bxy + cy^2$. When is it positive definite?

$$f(x, y) = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2$$

So $\boldsymbol{A}$ is positive definite when $a > 0$ and $ac - b^2 > 0$.

From those conditions, ***both eigenvalues are positive***.

(Hint: check the determinant and the trace)

It also shows that ***two pivots are also positive***.

(Hint: $\begin{bmatrix} a & b \\ b & c \end{bmatrix} \rightarrow \begin{bmatrix} a & b \\ 0 & (ac - b^2)/a \end{bmatrix}$)

However, it is not enough to require that the determinant of $\boldsymbol{A}$ *is positive*. For example, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ has positive determinant, but $\boldsymbol{A}$ is negative definite (负定).

The determinant test is applied not only to $\boldsymbol{A}$ itself, giving $ac - b^2 > 0$, but also to the 1 by 1 submatrix a in the upper left-hand corner.

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Both leading submatrices: $\boldsymbol{A}_1 = [a]$, $\boldsymbol{A}_2 = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ need to have positive determinant to ensure the positive definiteness of the matrix $\boldsymbol{A}$.

If $\boldsymbol{A} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 3 & -2 \\ -2 & -2 & 4 \end{bmatrix}$.

The 3 leading submatrices:

$$\boldsymbol{A}_1 = [2], \quad \boldsymbol{A}_2 = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}, \quad \boldsymbol{A}_3 = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 3 & -2 \\ -2 & -2 & 4 \end{bmatrix}.$$

Let $\boldsymbol{A}$ be a matrix of degree n, and let $\boldsymbol{A}_k$ be the *leading submatrix (upper left submatrices)* of degree k, consisting of the entries of $\boldsymbol{A}$ in the first k rows and first k columns.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & a_{1k+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & a_{2k+1} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & a_{kk+1} & \cdots & a_{kn} \\ a_{k+11} & a_{k+12} & \cdots & a_{k+1k} & a_{k+1k+1} & \cdots & a_{k+1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & a_{nk+1} & \cdots & a_{nn} \end{bmatrix}$$

$$\boldsymbol{A}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

- If $\boldsymbol{A} = \boldsymbol{LDU}$, then $\boldsymbol{A}_k = \boldsymbol{L}_k \boldsymbol{D}_k \boldsymbol{U}_k$, and $|\boldsymbol{A}_k| = d_1 d_2 \cdots d_k$.

Formula for pivots: $d_k = \dfrac{|\boldsymbol{A}_k|}{|\boldsymbol{A}_{k-1}|}$

Here is the main theorem on ***positive definiteness***, and a reasonably detailed proof.

Theorem 1 (Test for *positive definiteness*, Part I) Each of the following tests is a necessary and sufficient condition for the real symmetric matrix $\boldsymbol{A}$ to be ***positive definite***:

(正定性判别：以下是判定一个实对称矩阵$\boldsymbol{A}$正定的充要条件)

(I) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for all nonzero real vectors $\boldsymbol{x}$. (Definition)

(II) All the eigenvalues of $\boldsymbol{A}$ satisfy $\lambda_i > 0$.

(III) All the upper left submatrices $\boldsymbol{A}_k$ have positive determinants.

(IV) All the pivots (without row exchanges) satisfy $d_k > 0$.

Proof :

$$\begin{array}{ccc} \mathbf{I} & \leftrightarrow & \mathbf{II} \\ \uparrow & \searrow & \\ \mathbf{IV} & \leftarrow & \mathbf{III} \end{array}$$

Proof (I) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for all nonzero real vectors $\boldsymbol{x}$.

(II) All the eigenvalues of $\boldsymbol{A}$ satisfy $\lambda_i > 0$.

(I) ➔ (II): If $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, then $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\lambda\boldsymbol{x} = \lambda\|\boldsymbol{x}\|^2$.

A positive definite matrix has positive eigenvalues, since $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for $\boldsymbol{x} \neq \boldsymbol{0}$.

(II) ➔ (I): Since real symmetric matrices have a full set of orthonormal eigenvectors (denoted by $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$), any $\boldsymbol{x}$ is a combination $c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$. Then

$$\boldsymbol{A}\boldsymbol{x} = c_1\boldsymbol{A}\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{A}\boldsymbol{x}_n = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n.$$

Because of the orthogonality $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j = 0$ $(i \neq j)$, and the normalization $\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_i = 1$,

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} &= (c_1\boldsymbol{x}_1^{\mathrm{T}} + \cdots + c_n\boldsymbol{x}_n^{\mathrm{T}})(c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n)\\ &= c_1^2\lambda_1 + \cdots + c_n^2\lambda_n.\end{aligned}$$

If every $\lambda_i > 0$, then this equation shows that $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$.

Proof (I) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for all nonzero real vectors $\boldsymbol{x}$.

(III) All the upper left submatrices $\boldsymbol{A}_k$ have positive determinants.

(I) ➔ (III): The determinant of $\boldsymbol{A}$ is the product of the eigenvalues.

And if condition I holds, we already know that these eigenvalues are positive.

But we also have to deal with *every* upper left submatrix $\boldsymbol{A}_k$. The trick is to look at all nonzero vectors whose last $n-k$ components are zero:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix}\boldsymbol{x}_k^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_k & * \\ * & *\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_k \\ \boldsymbol{0}\end{bmatrix} = \boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{x}_k > 0.$$

Thus $\boldsymbol{A}_k$ is positive definite. Its eigenvalues (not the same λ_i) must be positive. Its determinant is their product, so all upper left determinants are positive.

That means: *If condition I holds, so does condition III.*

Proof (I) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for all nonzero real vectors $\boldsymbol{x}$.

(III) All the upper left submatrices $\boldsymbol{A}_k$ have positive determinants.

(IV) All the pivots (without row exchanges) satisfy $d_k > 0$.

(III) ➔ (IV): According to the formula for pivots, the kth pivot d_k is the ratio of $\det\boldsymbol{A}_k$ to $\det\boldsymbol{A}_{k-1}$.

If the determinants are all positive, so are the pivots.

(IV) ➔ (I): $\boldsymbol{A}$ is real symmetric, then $\boldsymbol{A} = \boldsymbol{LDL}^T$ where $\boldsymbol{L}$ is lower triangular, with 1's on the diagonal, and $\boldsymbol{D} = \mathrm{diag}(d_1, d_2, \cdots, d_n)$.

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T(\boldsymbol{LDL}^T)\boldsymbol{x} = (\boldsymbol{L}^T\boldsymbol{x})^T\boldsymbol{D}(\boldsymbol{L}^T\boldsymbol{x})$$

Let $\boldsymbol{y} = \boldsymbol{L}^T\boldsymbol{x} = [y_1 \quad y_2 \quad \cdots \quad y_n]^T$, then $\boldsymbol{x} \neq \boldsymbol{0} \Longrightarrow \boldsymbol{y} \neq \boldsymbol{0}$ since $\boldsymbol{L}$ is invertible.

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{D}\boldsymbol{y} = d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2 > 0, \forall \boldsymbol{x} \neq \boldsymbol{0}.$$

We illustrate the idea by giving an example.

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

Example 2 $\boldsymbol{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ has positive pivots 2, $\frac{3}{2}$, and $\frac{4}{3}$.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} = \boldsymbol{LDL}^{\mathrm{T}}.$$

If $\boldsymbol{x} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$, then $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{x} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} u - \frac{1}{2}v \\ v - \frac{2}{3}w \\ w \end{bmatrix}$.

So $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{LDL}^{\mathrm{T}}\boldsymbol{x} = \left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{x}\right)^{\mathrm{T}}\boldsymbol{D}\left(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{x}\right)$

$$= 2\left(u - \tfrac{1}{2}v\right)^2 + \tfrac{3}{2}\left(v - \tfrac{2}{3}w\right)^2 + \tfrac{4}{3}w^2 .$$

Those positive pivots in $\boldsymbol{D}$ multiply perfect squares to make $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ positive.

Example 2 Test the positive definiteness of the matrix

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

Each of the following tests is enough by itself. (选用任何一种即可)

(1) **Pivot test:** $\boldsymbol{A}$ has positive pivots $d_1 = 2, d_2 = \frac{3}{2}$, and $d_3 = \frac{4}{3}$.

(2) **Determinant test:** Each of the upper left submatrices has positive determinant: $|\boldsymbol{A}_1| = 2, |\boldsymbol{A}_2| = 3$, and $|\boldsymbol{A}_3| = |\boldsymbol{A}| = 4$.

(We can also see that $d_k = |\boldsymbol{A}_k|/|\boldsymbol{A}_{k-1}|$.)

(3) **Eigenvalue test:** The eigenvalues of $\boldsymbol{A}$ are $\lambda_1 = 2 - \sqrt{2}$, $\lambda_2 = 2$, and $\lambda_3 = 2 + \sqrt{2}$.

So *by any means*, we can conclude that $\boldsymbol{A}$ is positive definite.

Theorem 1 (Test for *positive definiteness*, Part II) Each of the following tests is a necessary and sufficient condition for the real symmetric matrix $\boldsymbol{A}$ to be ***positive definite***:

(V) There is a matrix $\boldsymbol{R}$ with independent columns such that $\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$.

Proof. (V) ➔ (I): *The key is to recognize*

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{x} = (\boldsymbol{R}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{R}\boldsymbol{x}) = \|\boldsymbol{R}\boldsymbol{x}\|^2.$$

This squared length $\|\boldsymbol{R}\boldsymbol{x}\|^2$ is positive (unless $\boldsymbol{x} = \boldsymbol{0}$), because $\boldsymbol{R}$ has independent columns. (If $\boldsymbol{x}$ is nonzero then $\boldsymbol{R}\boldsymbol{x}$ is nonzero.)

Thus $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{x} > 0$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$ is positive definite.

(I) ➔ (V): Since $\boldsymbol{A}$ is real symmetric, $\boldsymbol{A}$ can be decomposed into

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}} = (\boldsymbol{Q}\sqrt{\boldsymbol{\Lambda}})(\sqrt{\boldsymbol{\Lambda}}\boldsymbol{Q}^{\mathrm{T}}),$$

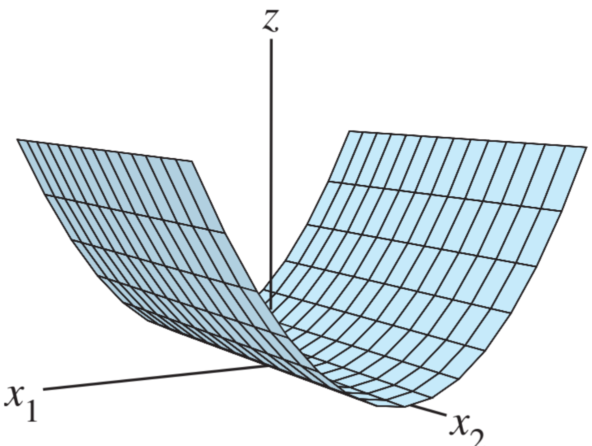
So take $\boldsymbol{R} = \sqrt{\boldsymbol{\Lambda}}\boldsymbol{Q}^{\mathrm{T}}$.

(Another choice is $\boldsymbol{R} = \boldsymbol{Q}\sqrt{\boldsymbol{\Lambda}}\boldsymbol{Q}^{\mathrm{T}}$, the ***symmetric positive definite square root*** of $\boldsymbol{A}$.)

II. Positive Semidefinite matrices (半正定矩阵)

Definition 2 A quadratic form $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is:

- **positive definite** if $f(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$;
- **negative definite** if $f(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \neq \mathbf{0}$;
- **indefinite** if $f(\boldsymbol{x})$ assumes both positive and negative values (既有正值又有负值);
- **positive semidefinite** (半正定) if $f(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$;
- **negative semidefinite** if $f(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x}$.



$z = 3x_1^2$

How to test the positive *semi*definiteness?

The main point is to see the ***analogies*** with the positive definite case.

(半正定性的判定: 类比正定性)

Theorem 2 Each of the following tests is a necessary and sufficient condition for a symmetric matrix $\boldsymbol{A}$ to be ***positive semidefinite***:

(I') $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq 0$ for all vectors $\boldsymbol{x}$. (This defines positive semidefinite)

(II') All the eigenvalues of $\boldsymbol{A}$ satisfy $\lambda_i \geq 0$.

(III') No principal submatrices have negative determinants.

(IV') No pivots are negative.

(V') There is a matrix $\boldsymbol{R}$, possibly with dependent columns, such that $\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$.

Remark The diagonalization $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$ leads to $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}$, where $\boldsymbol{y} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{x}$.

If $\boldsymbol{A}$ has rank r, there are r nonzero λ's and r perfect squares in $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2$. (We will come back to discuss this later)

Principal Submatrix

Let $\boldsymbol{A}$ be an $n \times n$ matrix. A $k \times k$ submatrix of $\boldsymbol{A}$ formed by deleting $n - k$ rows of $\boldsymbol{A}$, and the same $n - k$ columns of $\boldsymbol{A}$, is called **principal submatrix (主子矩阵)** of $\boldsymbol{A}$. The determinant of a principal submatrix of $\boldsymbol{A}$ is called a **principal minor (主子式)** of $\boldsymbol{A}$.

(Note that the definition does not specify which $n - k$ rows and columns to delete, only that *their indices must be the same*.)

For example, $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

First order principal submatrices: $[a_{11}], [a_{22}], [a_{33}]$.

Second order principal submatrices: $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}, \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$.

Third order principal submatrix: $\boldsymbol{A}$.

(III') No principal submatrices have negative determinants.

Note Condition (III') applies to all the principal submatrices, not only those in the upper left-hand corner (leading principal submatrices).

(判定半正定性时, 不仅要检查左上角各阶子矩阵的行列式, 即<u>**顺序主子式**</u>, 而且检查所有各阶主子矩阵的行列式即<u>**主子式**</u>)

Otherwise, we could not distinguish between two matrices whose upper left determinants were all zero:

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ is positive semidefinite;}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \text{ is negative semidefinite.}$$

Example 3 Show the positive *semi*definiteness of the following matrix by all five tests. (任选一种都是充分的)

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 0 & -2 & 4 \end{bmatrix},$$

- (I') $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = x_1^2 + 2x_2^2 + 4x_3^2 + 2x_1x_2 - 4x_2x_3 = (x_1 + x_2)^2 + (x_2 - 2x_3)^2 \geq 0$ (zero if $-x_1 = x_2 = 2x_3$)
- (II') The eigenvalues are $\lambda_1 = 0,\ \lambda_2 = (7 - \sqrt{13})/2, \lambda_3 = (7 + \sqrt{13})/2$ (a zero eigenvalue).
- (III') All principal submatrices

$$1, \quad 2, \quad 4$$

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1, \qquad \begin{vmatrix} 1 & 0 \\ 0 & 4 \end{vmatrix} = 4, \qquad \begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix} = 4,$$

$$|\boldsymbol{A}| = 0.$$

Example 3 Show the positive *semi*definiteness of the following matrix by all five tests. (任选一种都是充分的)

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 1 & 2 & -2\\ 0 & -2 & 4\end{bmatrix},$$

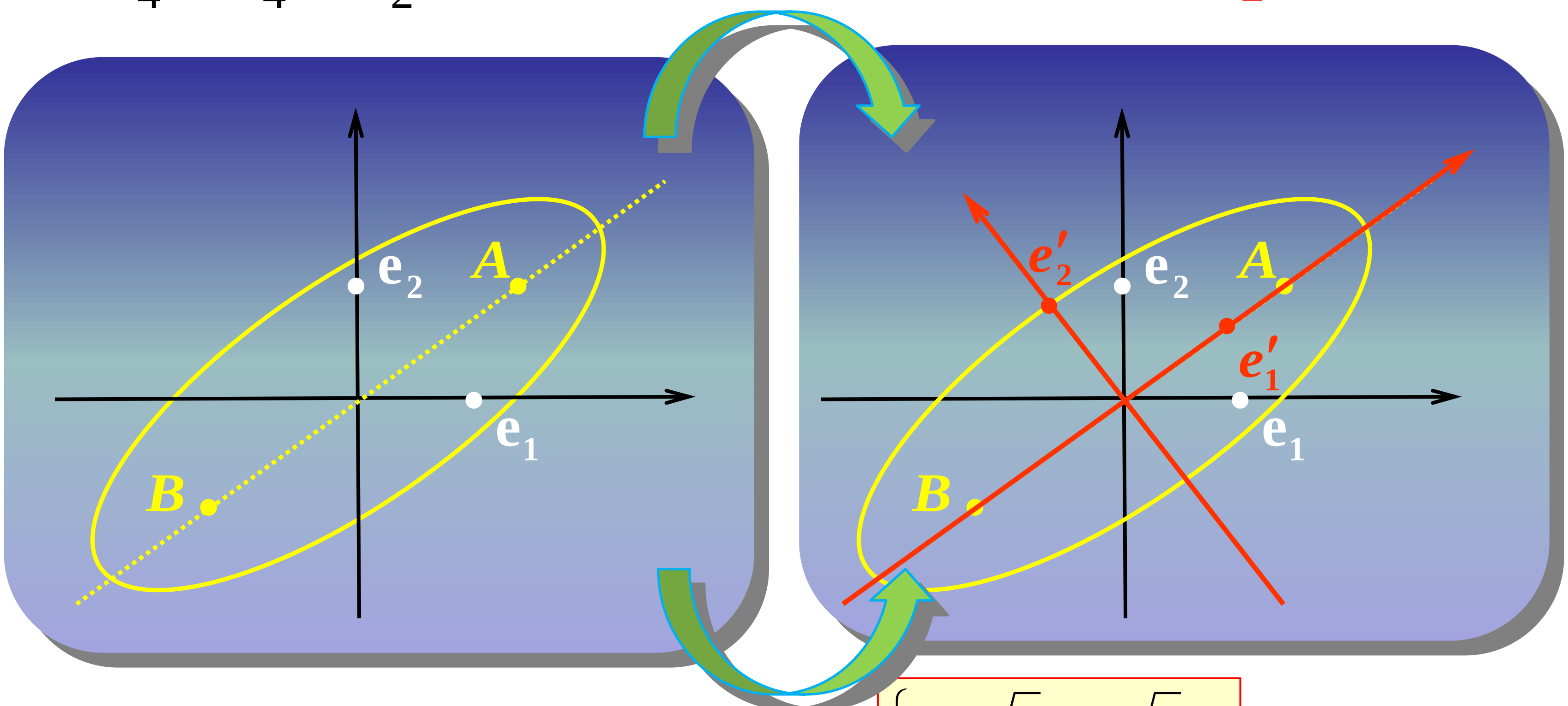
- (IV’) $\boldsymbol{A}\rightarrow\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & -2\\ 0 & -2 & 4\end{bmatrix}\rightarrow\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & -2\\ 0 & 0 & 0\end{bmatrix}$ (*missing* pivot)
- (V’) $\boldsymbol{A}=\boldsymbol{R}^T\boldsymbol{R}$ with dependent columns in the matrix $\boldsymbol{R}$:

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

Therefore, *by any means*, we can conclude that $\boldsymbol{A}$ is positive semidefinite.

III. The Principal Axes Theorem (主轴定理)

$$\frac{5}{4}x^2+\frac{5}{4}y^2-\frac{3}{2}xy=1$$ “消除交叉项” $$\frac{x'^2}{2}+2y'^2=1$$



$$\frac{1}{2}\left(\frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y\right)^2+2\left(\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y\right)^2=1$$

$$\boxed{\begin{cases} x=\dfrac{\sqrt{2}}{2}x'-\dfrac{\sqrt{2}}{2}y' \\ y=\dfrac{\sqrt{2}}{2}x'+\dfrac{\sqrt{2}}{2}y' \end{cases}}$$

“改斜归正”

——变量替换

Change of Variable (变量替换) in a Quadratic Form

If $\boldsymbol{x}$ represents a variable vector in $\mathbf{R}^n$, then a **change of variable** is an equation of the form

$$\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}, \text{ or equivalently, } \boldsymbol{y} = \boldsymbol{C}^{-1}\boldsymbol{x} \tag{1}$$

where $\boldsymbol{C}$ is an invertible matrix and $\boldsymbol{y}$ is a new variable vector in $\mathbf{R}^n$.

Here $\boldsymbol{y}$ is the coordinate vector of $\boldsymbol{x}$ relative to the basis of $\mathbf{R}^n$ determined by the columns of $\boldsymbol{C}$.

If the change of variable (1) is made in a quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, then

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y} \tag{2}$$

and the new matrix of the quadratic form is $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$.

Definition 3 For an invertible matrix $\boldsymbol{C}$, the linear transformation $\boldsymbol{A} \to \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ is called a **congruence transformation (合同变换)**, which transforms the vector $\boldsymbol{y}$ to the vector $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, and the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ to the quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$.

When A is real symmetric: $Q^{-1}AQ = \Lambda = Q^{\mathrm{T}}AQ$

$$Q = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}, \ \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Q: orthogonal matrix

$x_1, \dots, x_n$: *orthonormal eigenvectors*

$\lambda_1, \dots, \lambda_n$: *real eigenvalues*

Since A in the quadratic form is real symmetric, the Spectral Theorem guarantees the following theorem.

Theorem 3 (*The Principal Axes Theorem*, 主轴定理)
Let A be an $n \times n$ real symmetric matrix. Then there is an orthogonal change of variable, $x = Qy$ (i.e., Q is an *orthogonal* matrix), that transforms the quadratic form $x^{\mathrm{T}}Ax$ into a quadratic form $y^{\mathrm{T}}\Lambda y$ with no cross-product term (不含交叉乘积项) (i.e., Λ is a diagonal matrix).

$$x^{\mathrm{T}}Ax = y^{\mathrm{T}}\Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \text{ (二次型的标准形)},$$

where $\lambda_1, \cdots, \lambda_n$ are eigenvalues of A, and their orthonormal eigenvectors go into the columns of Q. $(Q^{\mathrm{T}}AQ = \Lambda)$

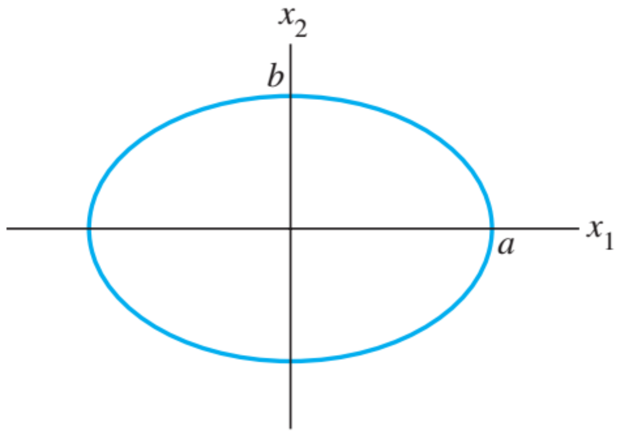
A Geometric View of Principal Axes

It can be shown that the set of all $\boldsymbol{x}$ in $\mathbf{R}^2$ that satisfy

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = c \qquad (c \text{ is a constant})$$

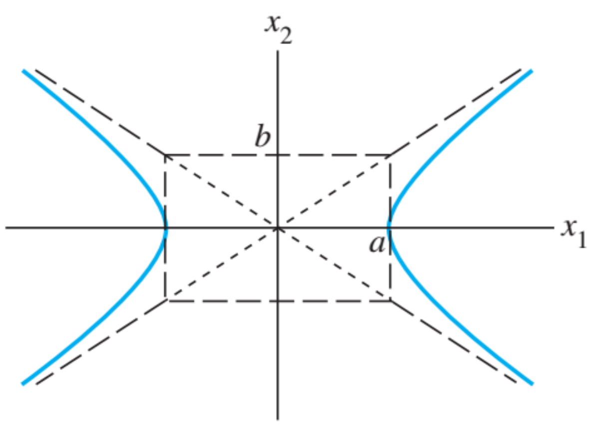
either corresponds to an ellipse (or circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.

If $\boldsymbol{A}$ is a diagonal matrix, the graph is in *standard position*, such as in the figure below.



$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \quad a > b > 0$$

ellipse

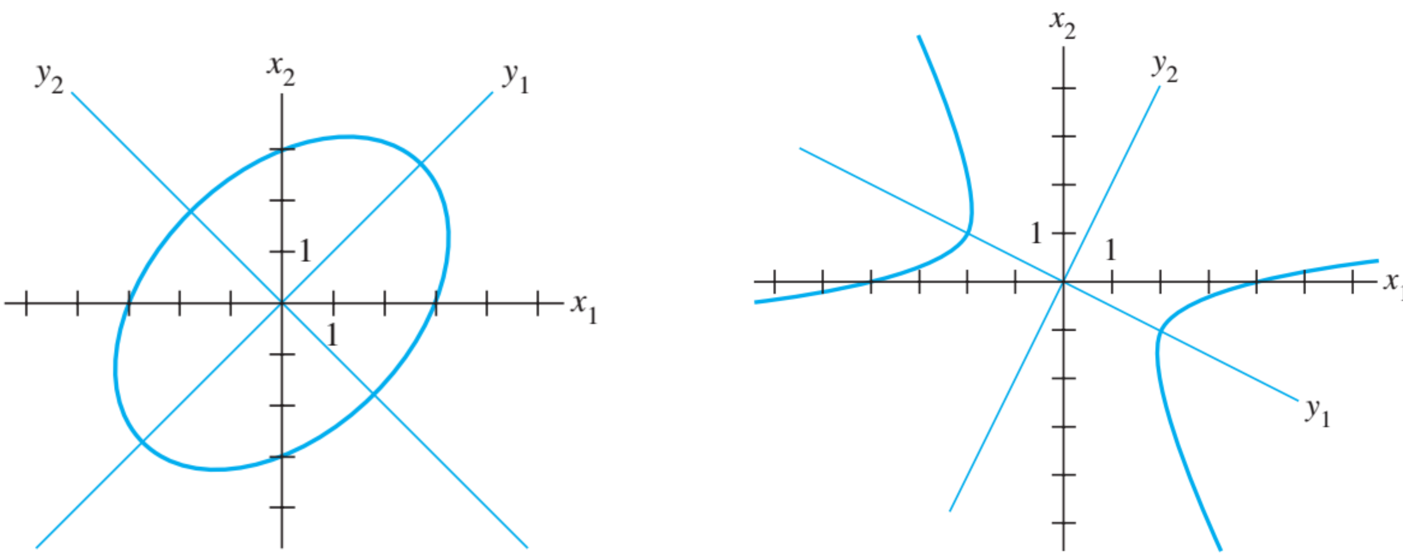


$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \quad a > b > 0$$

hyperbola

A Geometric View of Principal Axes

If ***A*** is not a diagonal matrix, the graph is rotated out of standard position, as in the figure below.



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

(b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

Finding the principal axes (determined by the eigenvectors of ***A***) amounts to finding a new coordinate system with respect to which the graph is in standard position.

Example 4(a) Let $\boldsymbol{A} = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$, and $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 5x_1^2 - 4x_1x_2 + 5x_2^2$.

The eigenvalues of $\boldsymbol{A}$ are 3 and 7, with corresponding unit eigenvectors

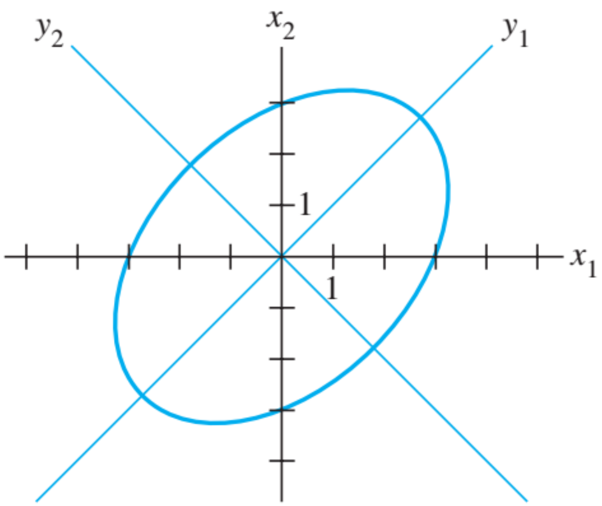
$$\boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Let $\boldsymbol{Q} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$.

Then $\boldsymbol{Q}$ orthogonally diagonalizes $\boldsymbol{A}$,

so the change of variable $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ produces

the quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = 3y_1^2 + 7y_2^2$.



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

Example 4(b) Let $\boldsymbol{A} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$, and $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = x_1^2 - 8x_1x_2 - 5x_2^2$.

The eigenvalues of $\boldsymbol{A}$ are 3 and -7, with corresponding unit

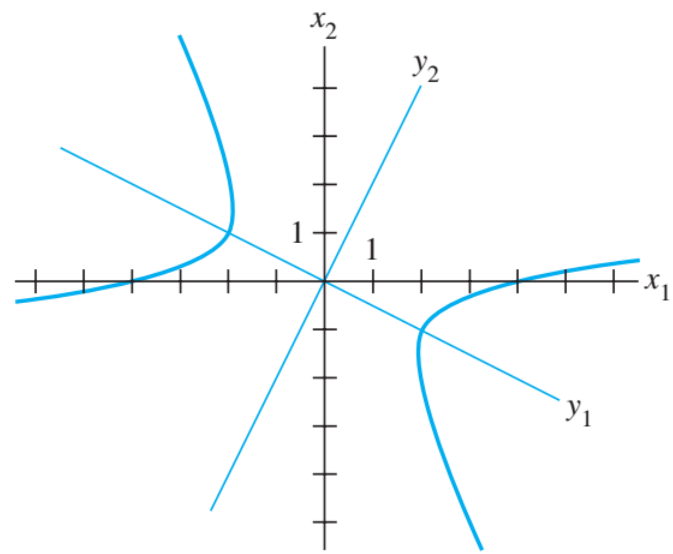
eigenvectors $\boldsymbol{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$.

Let $\boldsymbol{Q} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2] = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$.

Then $\boldsymbol{Q}$ orthogonally diagonalizes $\boldsymbol{A}$,

so the change of variable $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ produces

the quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = 3y_1^2 - 7y_2^2$.



(b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

The Principal Axes Theorem in n dimensions

Example 5 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 2x_1^2 + 4x_1x_2 - 4x_1x_3 + 5x_2^2 - 8x_2x_3 + 5x_3^2$

Converting this quadratic form into a standard form is just the same as:

(See Section 5.5 Example 2) Let $\boldsymbol{A} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$.

Find an orthogonal matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ is a diagonal matrix.

$$\boldsymbol{Q} = [\gamma_1, \gamma_2, \gamma_3] = \begin{bmatrix} -2\sqrt{5}/5 & 2\sqrt{5}/15 & 1/3 \\ \sqrt{5}/5 & 4\sqrt{5}/15 & 2/3 \\ 0 & \sqrt{5}/3 & -2/3 \end{bmatrix}$$

and

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 10 \end{bmatrix}.$$

So

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y} = y_1^2 + y_2^2 + 10y_3^2$$

Application: The quadric surface (二次曲面):

$$2x_1^2+4x_1x_2-4x_1x_3+5x_2^2-8x_2x_3+5x_3^2=1$$

has changed to

$$y_1^{\,2}+y_2^{\,2}+10y_3^{\,2}=1$$

with respect to a new basis $\{\gamma_1, \gamma_2, \gamma_3\}$.

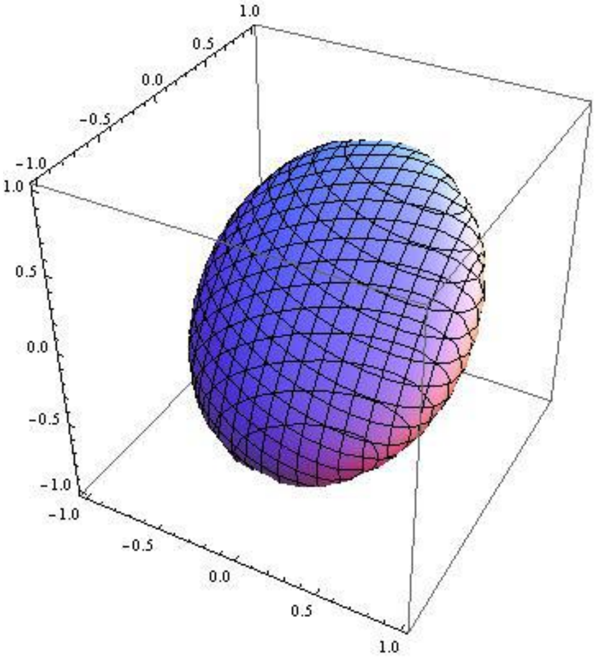
This is an equation of ellipsoid (椭球面), whose axes have half-lengths

$$\frac{1}{\sqrt{|\lambda_1|}}=1,\ \frac{1}{\sqrt{|\lambda_2|}}=1,\ \frac{1}{\sqrt{|\lambda_3|}}=\frac{1}{\sqrt{10}}.$$

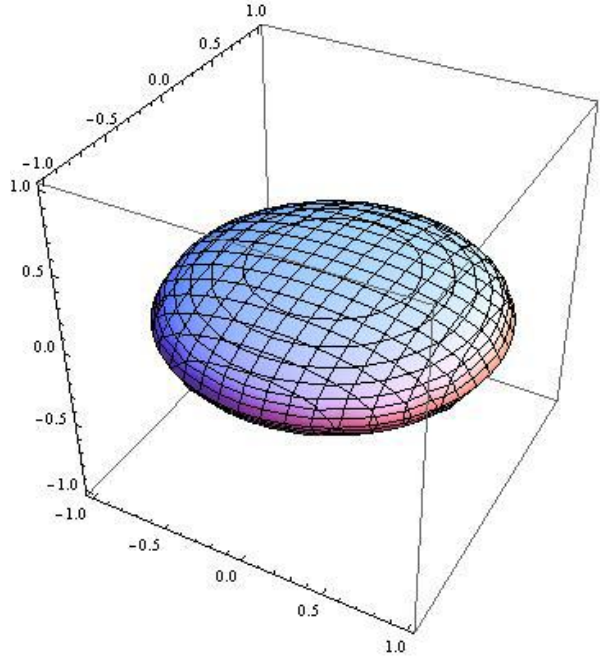
Remark Suppose $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$ ($\boldsymbol{A}$ is $n\times n$) with $\lambda_i>0$. Rotating $\boldsymbol{y}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{x}$ simplifies $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}=1=\lambda_1y_1^2+\cdots+\lambda_ny_n^2=1.$$

This is the equation of an ellipsoid. Its axes have *half*-lengths $\frac{1}{\sqrt{\lambda_1}},\ldots,\frac{1}{\sqrt{\lambda_n}}$ from the center. In the original $\boldsymbol{x}$-space they point along the eigenvectors of $\boldsymbol{A}$.



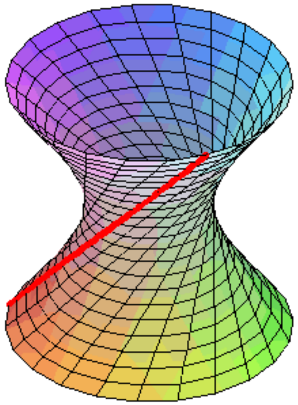
$$2x_1^2+4x_1x_2-4x_1x_3+5x_2^2-8x_2x_3+5x_3^2=1$$

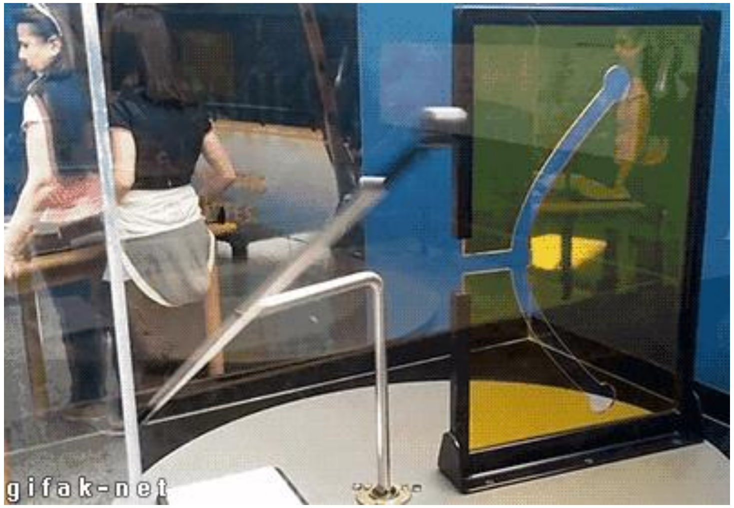


$$y_1^2+y_2^2+10y_3^2=1$$

特征值的符号决定了二次曲面的类型

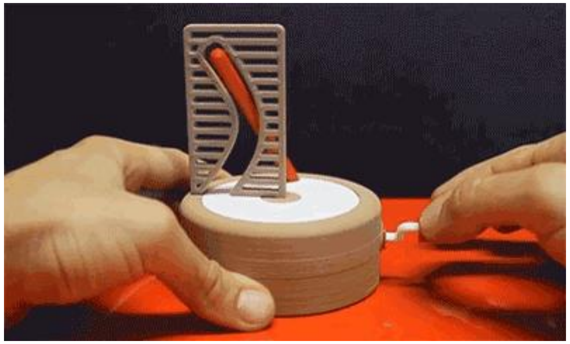
hyperboloid of one sheet (单叶双曲面)





Standard equation:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$



柯西指出：当方程是标准型时，二次曲面用二次项的符号来进行分类.

http://www.math.umn.edu/~rogness/quadrics/

http://mathinsight.org/quadric_surfaces

Quadratic equation (二次方程)

A quadratic equation in n variables $x_1, x_2, \dots, x_n$ is one of the form

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{x} + \alpha = 0,$$

where $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^{\mathrm{T}}$, $\boldsymbol{A}$ is an $n \times n$ symmetric matrix, $\boldsymbol{B}$ is a $1 \times n$ matrix, and α is a scalar.

The vector function $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is the quadratic form in n variables associated with the quadratic equation.

In the case of *three* unknowns, if

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \ \boldsymbol{A} = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \ \boldsymbol{B} = \begin{pmatrix} g & h & i \end{pmatrix},$$

then the quadratic equation is

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + \alpha = 0.$$

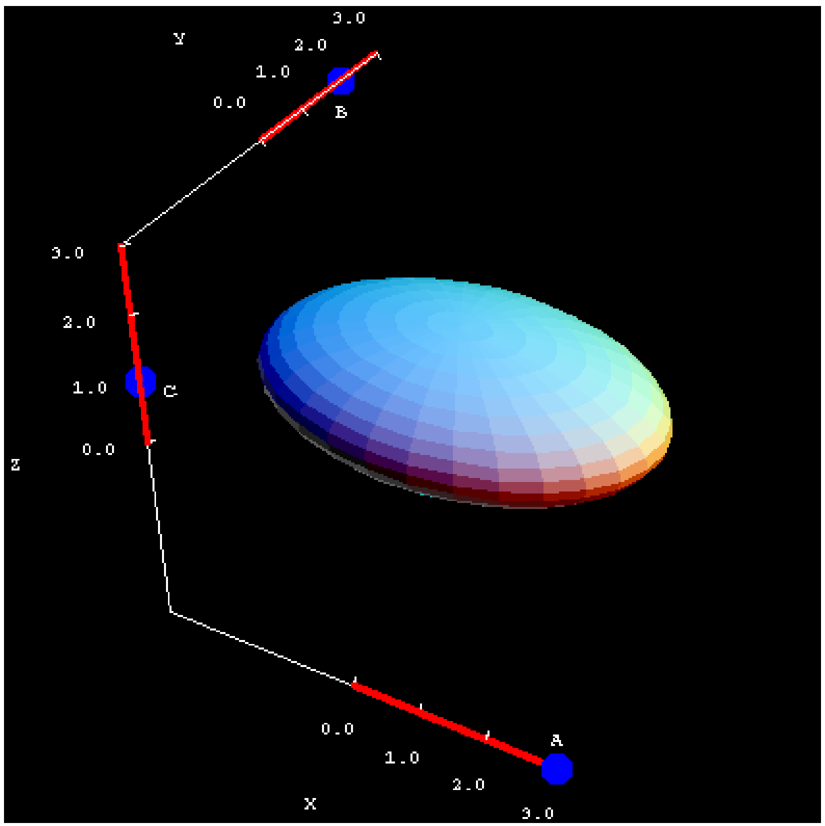
The graph of a quadratic equation in three variables is called a quadric surface (二次曲面).

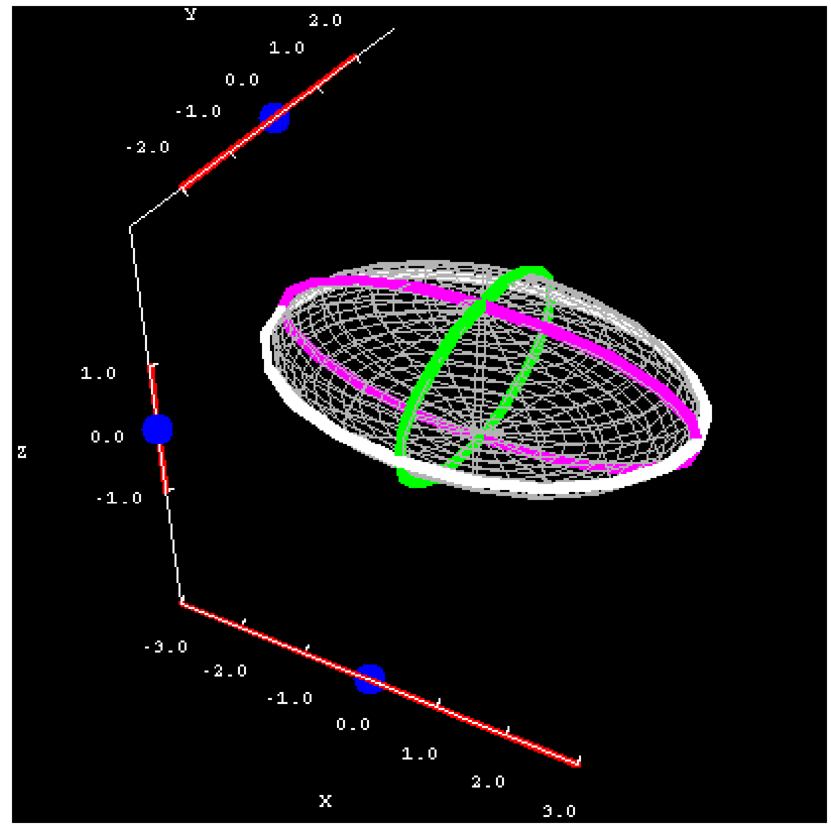
Some typical quadric surfaces (常见二次曲面分类)

1. Ellipsoid （椭球面）
2. Hyperboloid of one sheet（单叶双曲面）
3. Hyperboloid of two sheets（双叶双曲面）
4. Elliptic paraboloid（椭圆抛物面）
5. Hyperbolic paraboloid（双曲抛物面）
6. Double cone（圆锥面）

1. Ellipsoid （椭球面）

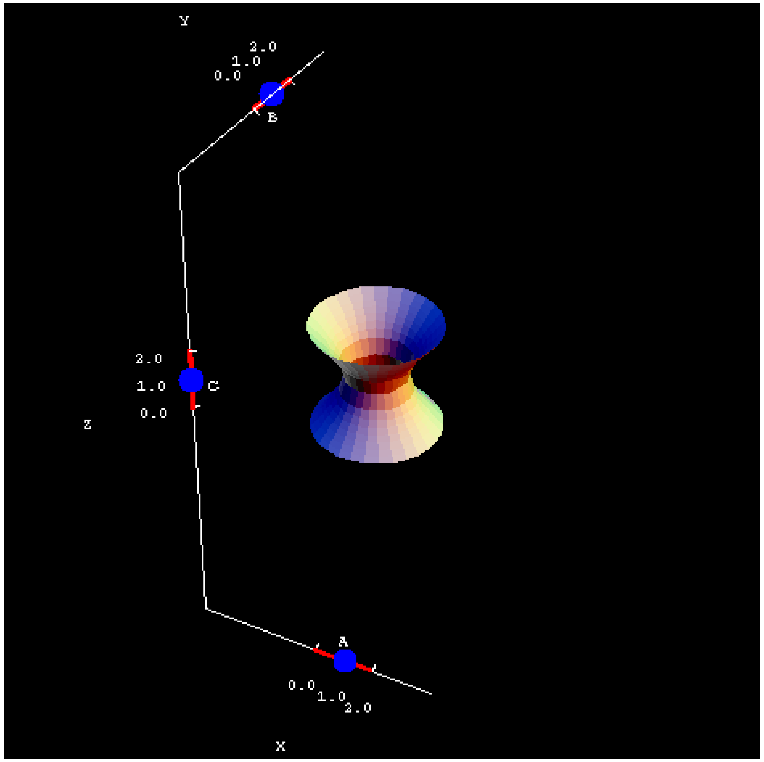
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \quad (a,b,c>0)$$

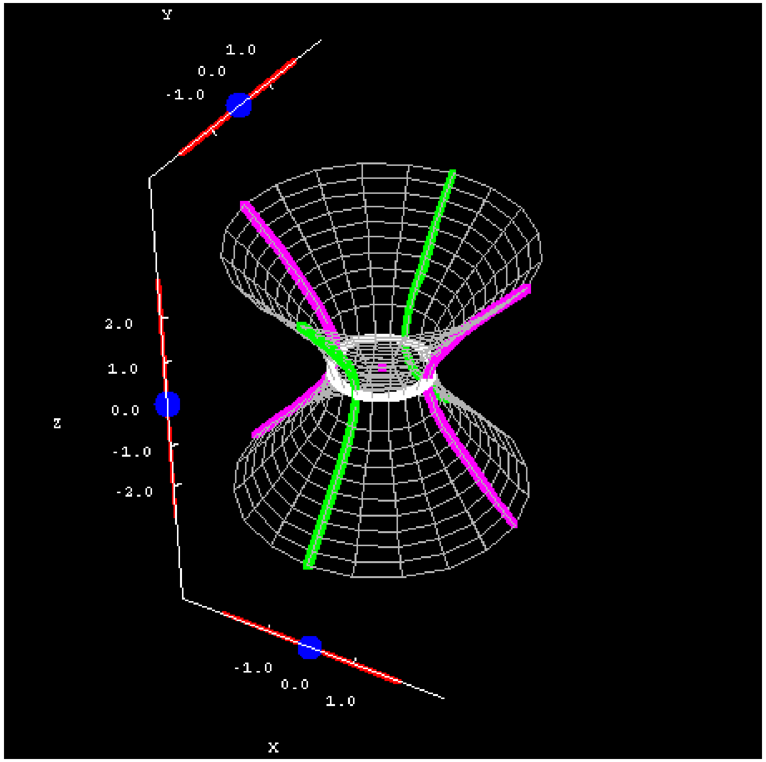




2. Hyperboloid of one sheet（单叶双曲面）

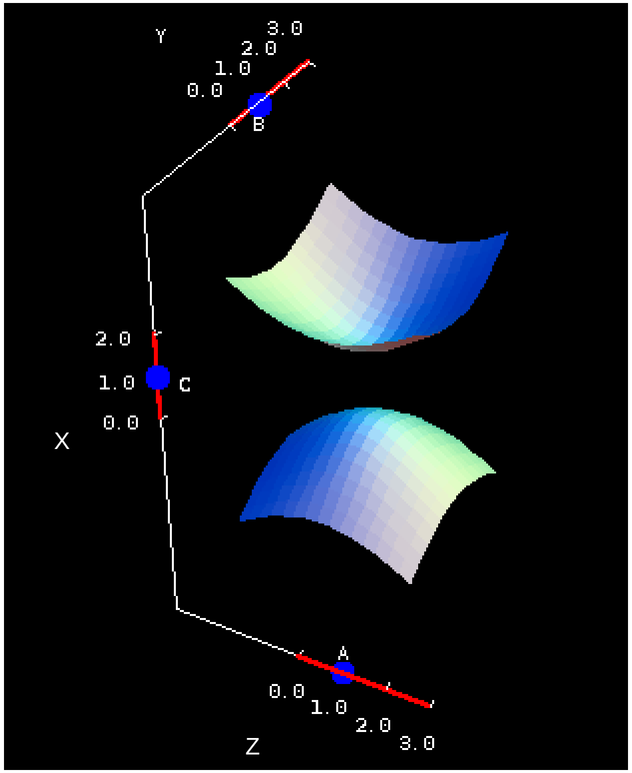
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,b,c>0)$$

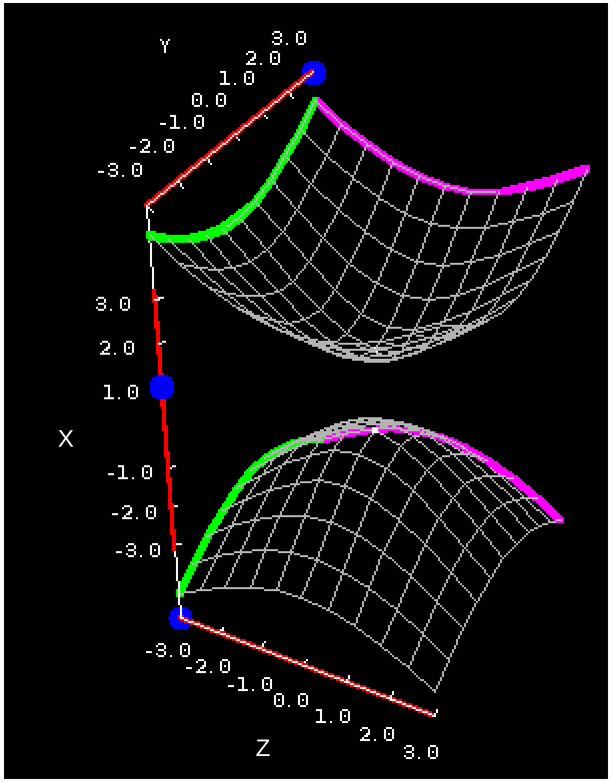




3. Hyperboloid of two sheets（双叶双曲面）

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,b,c>0)$$

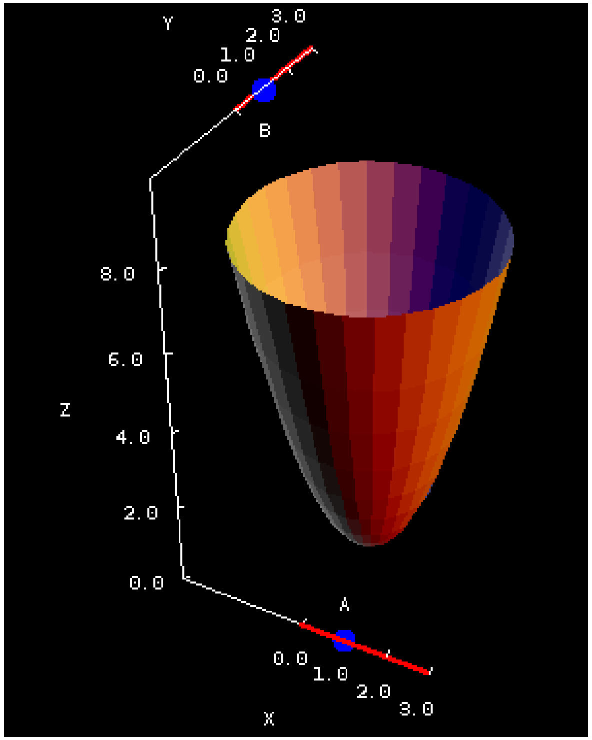


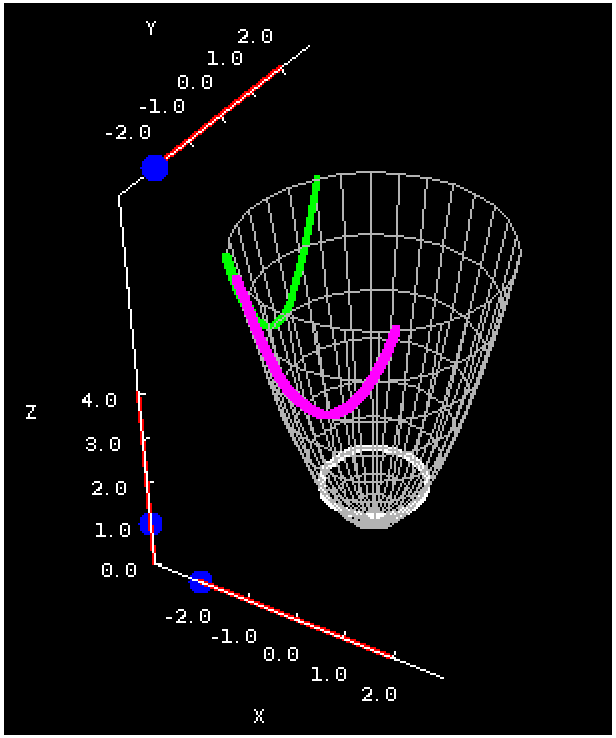


4. Elliptic paraboloid（椭圆抛物面）

$$\frac{x^2}{2p}+\frac{y^2}{2q}=z,$$

where p, q have the same sign.

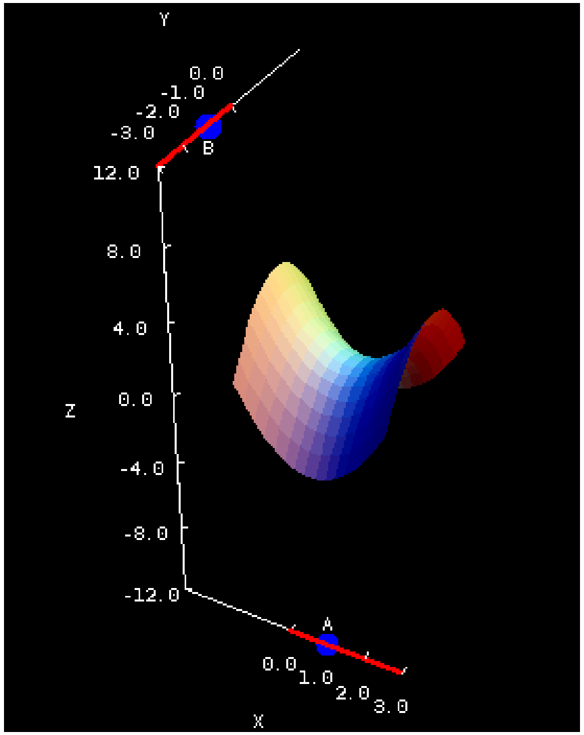


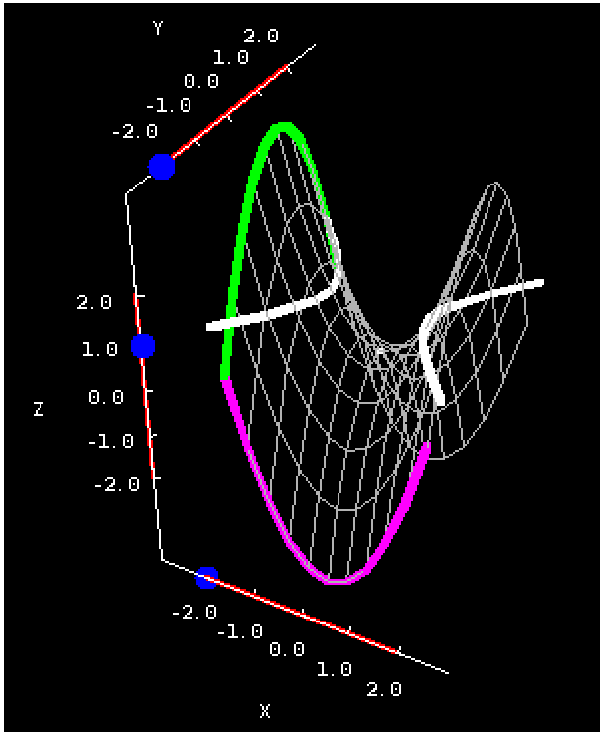


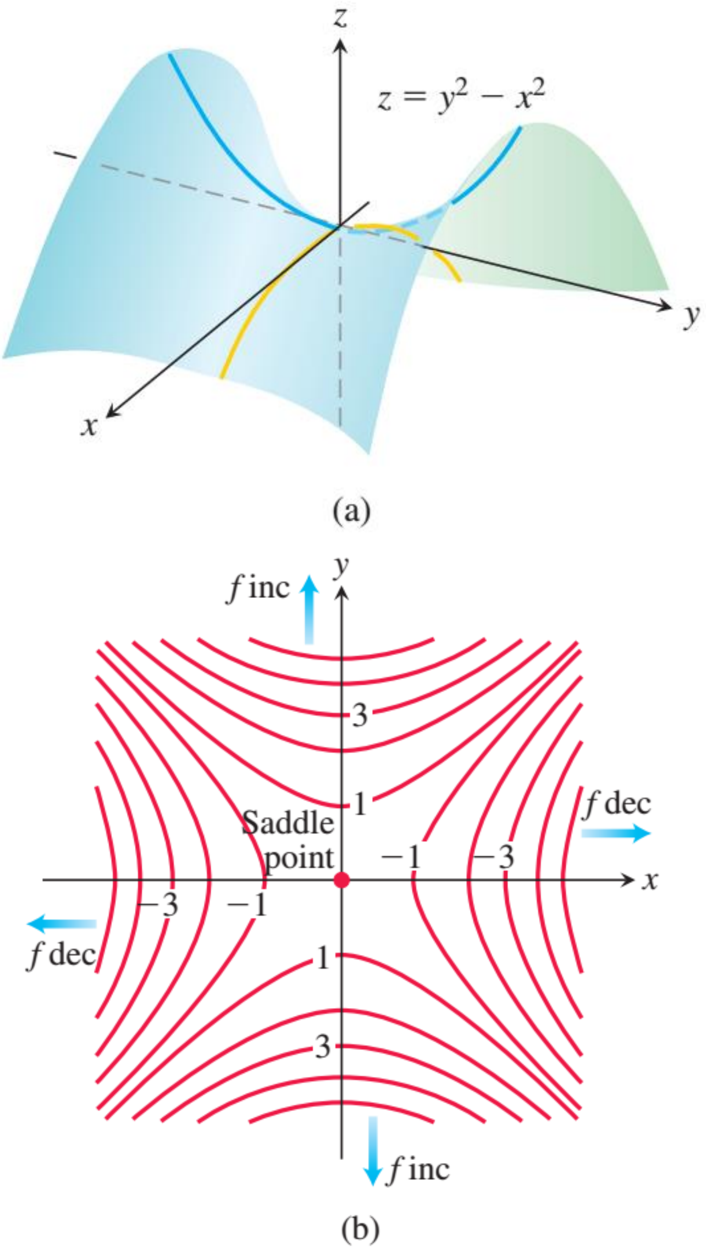
5. Hyperbolic paraboloid（双曲抛物面）

$$\frac{x^2}{2p} - \frac{y^2}{2q} = z,$$

where p, q have the same sign.

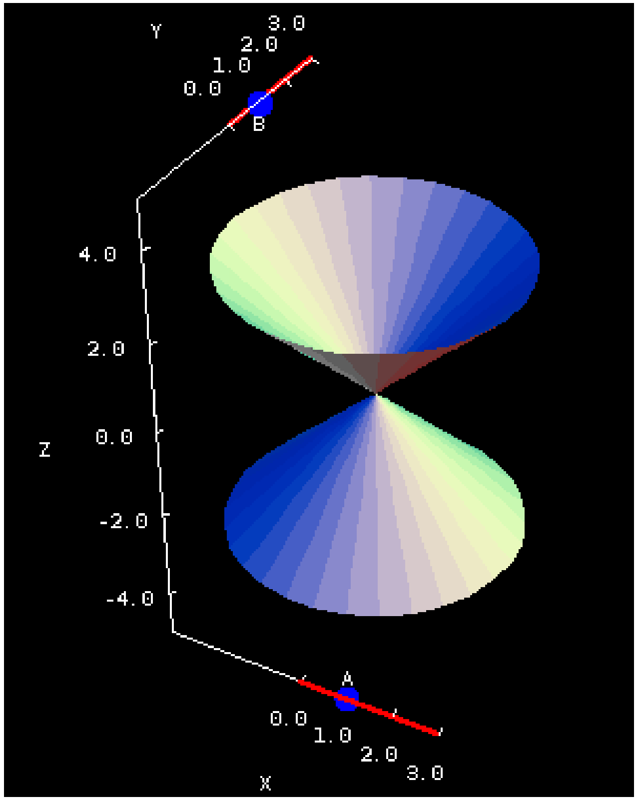


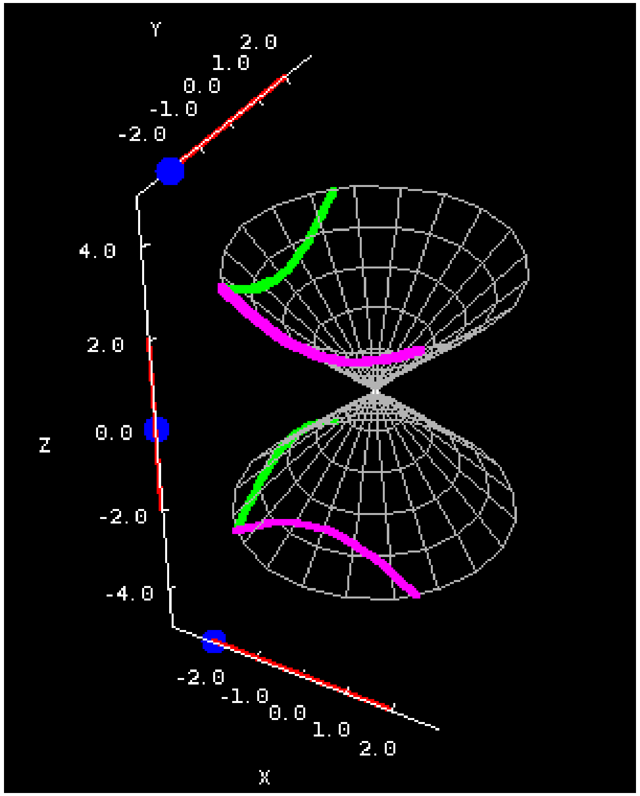



z
$z = y^2 - x^2$
y
x
(a)
y
f inc
3
1
Saddle
point
−1
−3
f dec
x
−3
−1
f dec
1
3
f inc
(b)

6. Double cone（圆锥面）

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \quad (a,b,c>0)$$





IV. The Law of Inertia (惯性定律)

Matrix A	Operations (变换)	Matrix B	Invariants (不变量)
$\boldsymbol{A}$ is any $m \times n$ matrix	Elementary operations	$\boldsymbol{B} = \boldsymbol{PAQ}$ (where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are invertible $m \times m$ and $n \times n$ matrices)	Rank
$\boldsymbol{A}$ is any $n \times n$ matrix	Similarity transformation (相似变换)	$\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{AM}$ (where $\boldsymbol{M}$ is an invertible $n \times n$ matrix)	Eigenvalues; Determinant; Trace; Rank
$\boldsymbol{A}$ is any real symmetric $n \times n$ matrix	Congruence transformation (合同变换)	$\boldsymbol{B} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ (where $\boldsymbol{C}$ is an invertible $n \times n$ matrix)	Symmetry; Rank; Number of positive eigenvalues, negative eigenvalues, and zero eigenvalues

Theorem 4 (*Sylvester's law of inertia*)

$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ has the same number of positive eigenvalues, negative eigenvalues, and zero eigenvalues as $\boldsymbol{A}$, where $\boldsymbol{C}$ is a *nonsingular* matrix.

[在化简成标准型时, 为何总是得到同样数目的正项和负项?

西尔维斯特 (James Joseph Sylvester, 英国数学家, 1814-1897) 给出了二次型的惯性定律, 但没证明.

该定律后被**雅可比** (Jacobi, Carl Gustav Jacob, 德国数学家, 1804-1851) 重新发现和证明.]

Remark.

For any real symmetric matrix $\boldsymbol{A}$, ***the signs of the pivots agree with the signs of the eigenvalues***. The eigenvalue matrix $\boldsymbol{\Lambda}$ and the pivot matrix $\boldsymbol{D}$ have the same number of positive entries, negative entries, and zero entries.

(We will assume that $\boldsymbol{A}$ allows the symmetric factorization $\boldsymbol{A} = \boldsymbol{LDL}^T$ (without row exchanges).

By the law of inertia, $\boldsymbol{A}$ has the same number of positive eigenvalues as $\boldsymbol{D}$.

But the eigenvalues of $\boldsymbol{D}$ are just its diagonal entries (the pivots). Thus the number of positive pivots matches the number of positive eigenvalues of $\boldsymbol{A}$.)

That is both beautiful and practical. (*brings together pivots and eigenvalues; uses pivots to locate eigenvalues*)

Example 6 (a) Suppose $\boldsymbol{A} = \boldsymbol{I}$.

Then $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ is positive definite.

Both $\boldsymbol{I}$ and $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$ have n positive eigenvalues, confirming the law of inertia.

Example 6 (b) If $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$,

then $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ has a negative determinant:

$$\left|\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\right| = \left|\boldsymbol{C}^{\mathrm{T}}\right| \cdot |\boldsymbol{A}| \cdot |\boldsymbol{C}| = -|\boldsymbol{C}|^2 < 0,$$

therefore $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ must have one positive and one negative eigenvalue, like $\boldsymbol{A}$.

Key words:

Tests for Positive Definiteness;
Semidefinite Matrices;
The Principal Axes Theorem;
The Law of Inertia

Homework

See Blackboard