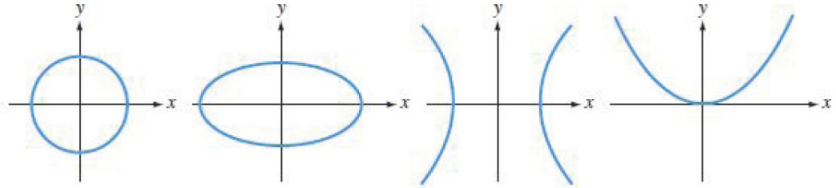
Slide 26-Quadratic Form

MAT2040 Linear Algebra

Introduction

Recall: $x^2 + y^2 = r^2$ – – – – – – – – – – – – – – – – –(*circle*)
$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$ – – – – – – – – – – – – – – – – –(*ellipse*)
$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = \pm 1$ – – – – – – – – – – – – – – – – –(*hyperbola*)
$x^2 = \alpha y, \quad \text{or}, \quad y^2 = \alpha x$ – – – – – – – – – – – – – – – – –(*parabola*)



(i) Circle (ii) Ellipse (iii) Hyperbola (iv) Parabola

Question

How to classify the type for the quadratic equation with two unknowns:

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0?$$

Definition 26.1 (Quadratic Equation with two unknowns) A quadratic equation in two unknowns is an equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 - - - - - - - - - - - - - - - (*)$$

(*) can be written in the form

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + f = 0$$

Let

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

then (*) can be written as

$$\mathbf{x}^T A\mathbf{x} + [d, e]\mathbf{x} + f = 0$$

The term $\mathbf{x}^T A\mathbf{x}$ is called the quadratic form associated with quadratic equation (*). The graph corresponding to (*) is called the **conic section**.

The classification of the conic section of (*) is completely solved. Here I only provide one example to illustrate the classification idea.

Example 26.2

$$3x^2 + 2xy + 3y^2 + 8\sqrt{2}y - 4 = 0$$

Here $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ is a real symmetric matrix. Now we can take $Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ to be an orthogonal matrix such that

$$Q^T AQ = diag(2, 4).$$

Let $\begin{bmatrix} x \\ y \end{bmatrix} = Q \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$(changing the coordinate system)

Then

$$2(\hat{x})^2 + 4(\hat{y})^2 - 8\hat{x} + 8\hat{y} = 4$$

which equivalent to

$$2(\hat{x} - 2)^2 + 4(\hat{y} + 1)^2 = 16$$

i.e.,

$$\frac{(\hat{x} - 2)^2}{8} + \frac{(\hat{y} + 1)^2}{4} = 1$$

which is an ellipse.

Observation: The quadratic term $3x^2 + 2xy + 3y^2$ determines the type of conic section of this quadratic equation. The quadratic term plays important role in determining the type of the conic section.

Quadratic form with n variables

Consider the quadratic form $f(\mathbf{x}) = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}x_j)x_i$, where $A = (a_{ij})_{n\times n}$ is a real matrix, $\mathbf{x} = (x_i)_{n\times 1}$ is a real column vector.

Since $\mathbf{x}^T A\mathbf{x} = (\mathbf{x}^T A\mathbf{x})^T = \mathbf{x}^T A^T\mathbf{x}$. Thus, $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \frac{A+A^T}{2}\mathbf{x}$.
Only need to discuss the symmetric real matrix A.

In the following, we focus on the quadratic term

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

where A is a real symmetric matrix.

Definition 26.3 (**Definite quadratic form and definite matrix**) Let $\mathbf{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ be symmetric, then

(1) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **positive definite** if $f(\mathbf{x}) > 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **positive definite matrix**.

(2) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **positive semidefinite** if $f(\mathbf{x}) \geq 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **positive semidefinite matrix**.

(3) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **indefinite** if $f(\mathbf{x})$ takes different signs.

Similarly, the negative definite and negative semidefinite can defined as follows:

(4) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **negative definite** if $f(\mathbf{x}) < 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **negative definite matrix**.

(5) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **negative semidefinite** if $f(\mathbf{x}) \geq 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **negative semidefinite matrix**.

Example 26.4

(1)

$$f(\mathbf{x}) = (x, y, z) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 + 2y^2 + 3z^2 > 0$$

if $(x, y, z) \neq (0, 0, 0)$, thus A is positive definite.

(2)

$$f(\mathbf{x}) = (x, y, z) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 + 2y^2 \geq 0$$

and $f(0, 0, 1) = 0$, thus A is positive semidefinite.

(3)

$$f(\mathbf{x}) = (x, y, z) \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 - 2y^2$$

thus A is indefinite.

Theorem 26.5 Let $A \in \mathbb{R}^{n \times n}$ be symmetric, then A is positive definite if only if all eigenvalues are positive.

Proof. Since A is symmetric, by spectral theorem for real symmetric matrix, there exists an orthogonal matrix Q such that $Q^{-1}AQ = Q^T AQ = D$, where D is the diagonal matrix. Let $\hat{\mathbf{x}} = Q^T \mathbf{x}$ then $\mathbf{x} = Q\hat{\mathbf{x}}$ and $\mathbf{x}^T A\mathbf{x} = (Q\hat{\mathbf{x}})^T AQ\hat{\mathbf{x}} = \hat{\mathbf{x}}^T Q^T AQ\hat{\mathbf{x}} = \hat{\mathbf{x}}^T D\hat{\mathbf{x}}$. Since Q is invertible and $\hat{\mathbf{x}} = Q^T \mathbf{x}$, thus

$$\mathbf{x}^T A\mathbf{x} > 0,\ \forall \mathbf{x} \neq \mathbf{0} \Leftrightarrow \hat{\mathbf{x}}^T D\hat{\mathbf{x}} > 0,\ \forall \hat{\mathbf{x}} \neq \mathbf{0}$$

Thus, A is positive definite $\Leftrightarrow$ the entries in diagonal elements of D are all positive$\Leftrightarrow$ all eigenvalues of A are positive.

Remark.

1. Let $A \in \mathbb{R}^{n \times n}$ be symmetric, then A is negative definite if only if all eigenvalues are negative.
2. Let $A \in \mathbb{R}^{n \times n}$ be symmetric, then A is indefinite if only if eigenvalues have different signs.

Corollary 26.6 Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite, then $\det(A) = \prod_{i=1}^{n} \lambda_i > 0$ ($\lambda_1, \cdots, \lambda_n$ are eigenvalues of A) and hence is invertible.

Example 26.7 The quadratic form

$$f(x, y) = 2x^2 - 4xy + 5y^2 = [x, y] \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

The eigenvalue of

$$\begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$$

is 6 and 1. Thus it is positive definite.

Definition 26.8 (**Leading Principal Submatrix**) Given A, let A_r denoted the matrix formed by deleting the last $n-r$ rows and last $n-r$ columns. A_r is called the leading principle submatrix of A of order r.

Property 26.9 (**Property of Positive Definite Matrix**) If A ($A \in \mathbb{R}^{n\times n}$) is a symmetric positive definite matrix, then all the leading principal submatrices $A_1, A_2, \cdots, A_n$ of A are all positive definite matrices, and thus all leading principal submatrices have positive determinants.

Proof. From the fact that

$$\begin{aligned}&[x_1, x_2, \cdots, x_r]A_r[x_1, x_2, \cdots, x_r]^T\\ =&[x_1, x_2, \cdots, x_r, 0, \cdots, 0]A[x_1, x_2, \cdots, x_r, 0, \cdots, 0]^T > 0\end{aligned}$$

for any $r = 1, 2 \cdots, n$ and $[x_1, x_2, \cdots, x_r]^T \neq \mathbf{0}$.
Thus, $A_r, r = 1, 2 \cdots, n$ are positive definite, thus have positive determinants.

(**Definition for Principal Submatrix**) Given A, let A_s denoted the submatrix formed by using rows $i_1, \cdots, i_s$ and columns $i_1, \cdots, i_s$ of A, then A_s is called the principal submatrix with order s.

Corollary If A ($A \in \mathbb{R}^{n \times n}$) is a symmetric positive definite matrix, then all the principal submatrices of A are all positive definite matrices, and thus all principal submatrices have positive determinants. In particular, all diagonal entries of A $a_{ii} > 0, i = 1, \cdots, n$.

Proof. Let A_s denoted the submatrix formed by using rows $i_1, \cdots, i_s$ and columns $i_1, \cdots, i_s$ of A, then

$$[x_{i_1}, \cdots, x_{i_s}]A_s[x_{i_1}, \cdots, x_{i_s}]^T = \mathbf{y}^T A\mathbf{y} > 0$$

for any $[x_{i_1}, \cdots, x_{i_s}]^T \neq \mathbf{0}$, where $\mathbf{y} = [y_1, \cdots, y_n]^T$ and

$$y_i = \begin{cases} 0, & \text{if } i \notin \{i_1, \cdots, i_s\}, \\ x_i, & \text{if } i \in \{i_1, \cdots, i_s\}. \end{cases}$$

Example $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$, Leading principal submatrices

$A_1 = [2], A_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, A_4 = A$, all have the positive determinants.

Take $\{i_1, i_2\}$=$\{2,3\}$, principal submatrix is $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, which has positive determinant.

Take $\{i_1, i_2, i_3\}$=$\{1,2,4\}$, principal submatrix is $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, which has positive determinant.

Property 26.10 Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite matrix, then

$$A = LU$$

(LU factorization)
where

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \ddots & 0 \\ * & \ddots & \ddots & 0 \\ * & * & * & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & * & \cdots & * \\ 0 & u_{22} & \ddots & * \\ 0 & \ddots & \ddots & * \\ 0 & 0 & 0 & u_{nn} \end{bmatrix}$$

L is a unit lower triangular matrix, and U is a upper triangular matrix whose diagonal elements are positive. In particular, A can be row reduced into U only by using the row operation III, the determinants for the leading principle matrices will not change during the Gaussian elimination process, thus the pivot elements $u_{11}, u_{22}, \cdots, u_{nn}$ will all positive.

An illustration for the Gaussian elimination of a 4×4 matrix is provided in the following figure.

$$\underset{A}{\begin{pmatrix} a_{11} & x & x & x \\ x & a_{22} & x & x \\ x & x & a_{33} & x \\ x & x & x & a_{44} \end{pmatrix}} \xrightarrow{1} \underset{A^{(1)}}{\begin{pmatrix} a_{11} & x & x & x \\ 0 & a_{22}^{(1)} & x & x \\ 0 & x & a_{33}^{(1)} & x \\ 0 & x & x & a_{44}^{(1)} \end{pmatrix}} \xrightarrow{2} \underset{A^{(2)}}{\begin{pmatrix} a_{11} & x & x & x \\ 0 & a_{22}^{(1)} & x & x \\ 0 & 0 & a_{33}^{(2)} & x \\ 0 & 0 & x & a_{44}^{(2)} \end{pmatrix}} \xrightarrow{3} \underset{A^{(3)} = U}{\begin{pmatrix} a_{11} & x & x & x \\ 0 & a_{22}^{(1)} & x & x \\ 0 & 0 & a_{33}^{(2)} & x \\ 0 & 0 & 0 & a_{44}^{(3)} \end{pmatrix}}$$

Figure: Elimination for 4×4 symmetric positive matrix, where it can be shown that $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, a_{44}^{(4)} > 0$

Example 26.11 Take the positive definite matrix

$$A = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix}$$

Then

$$L_2 L_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 3 & 4 \end{bmatrix}$$

$$L_3(L_2 L_1 A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

Thus

$$
\begin{aligned}
A &= L_1^{-1} L_2^{-1} L_3^{-1} \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{3} & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix} \\
&= LU
\end{aligned}
$$

Matrix U can be decomposed into

$$U = \begin{bmatrix} 4 & 2 & -2 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} = DU_1$$

L, U_1 are referred to the unit triangular matrices. It follows that $A = LDU_1$, which is called the LDU factorization of A (where L is a unit lower triangular matrix, D is a diagonal matrix with positive diagonal entries, U is a unit upper triangular matrix).

Remark. Indeed $U_1 = L^T$ for symmetric matrix A.

Theorem 26.12 Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite matrix, then

$$A = LDU \quad \text{(LDU factorization)}$$

where

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \ddots & 0 \\ * & \ddots & \ddots & 0 \\ * & * & * & 1 \end{bmatrix}, \quad D = \operatorname{diag}(u_{11}, u_{22}, \cdots, u_{nn})$$

$$U = \begin{bmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & * \\ 0 & \ddots & \ddots & * \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

L is a unit lower triangular matrix, and U is a unit upper triangular matrix, D is a diagonal matrix with positive diagonal entries. **Remark. Indeed $U = L^T$ and $A = LDL^T$.**

It will be an excise to check that the inverse of a unit lower triangular matrix is also a unit lower triangular matrix and the inverse of a unit upper triangular matrix is also a unit upper triangular matrix.

Theorem 26.13 Let A be a square matrix, if A has the LDU factorization, then the LDU factorization for A is unique.

Proof. Leave as an exercise.

Theorem 26.14 Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite matrix, then $A = LDL^T$ where L is a unit lower triangular matrix, D=diag$(d_1, \cdots, d_n)$ with $d_i > 0, i = 1, \cdots, n$.

Proof. By above theorem 25.14, $A = LDU$. And $A^T = U^T DL^T = LDU = A$. By using the uniqueness for LDU factorization, $U = L^T$ and $A = LDL^T$. Since A is positive definite, $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T LDL^T \mathbf{x} = (L^T\mathbf{x})^T D(L^T\mathbf{x}) > 0$ for any $\mathbf{x} \neq \mathbf{0}$, thus D must also positive definite, thus all elements in the diagonal of D are positive. Let $D^{\frac{1}{2}}$=diag$(\sqrt{d_1}, \cdots, \sqrt{d_n})$, then $A = LDL^T = LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T = L_1 L_1^T$, where $L_1 = LD^{\frac{1}{2}}$.

Remark If A is a symmetric positive definite matrix, then A can be factorized into a product LL^T, where L is lower triangular with positive diagonal elements. This is called the **Cholesky Decomposition**. And $A = LL^T = R^T R$, where $R = L^T$ is the upper triangular matrix.

Example 26.15

$$
\begin{aligned}
A &= \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 5 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \left[\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \right]^T \\
&= LL^T
\end{aligned}
$$

where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & \sqrt{3} \end{bmatrix}$$

Theorem 26.16 Let $A \in \mathbb{R}^{n \times n}$ and A is symmetric, then the following are equivalent

(1) A is positive definite.

(2) The leading principal submatrices $A_1, \cdots, A_n$ all have positive determinants.

(3) $A = LU$, where U is an upper triangular matrix with positive diagonal elements, L is a unit lower triangular matrix.

(4) $A = LDL^T$, where D is a diagonal matrix with positive diagonal elements, L is a unit lower triangular matrix.

(5) $A = LL^T$, where L is a lower triangular matrix with positive diagonal elements.

(6) $A = B^T B$ for some invertible matrix B.

Sketched Proof.

Since (1)$\Rightarrow$(2)$\Rightarrow$(3)$\Rightarrow$(4)$\Rightarrow$(5)$\Rightarrow$(6) has already been shown. Only show (6)$\Rightarrow$ (1).
$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T B^T B\mathbf{x} = \| B\mathbf{x} \|^2 > 0$ when $\mathbf{x} \neq \mathbf{0}$, this is because $\|B\mathbf{x}\| = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$.