Negative/positive (semi-)definite matrix and bordered Hessian matrix

1.A. Negative/positive (semi-)definite matrix

The definiteness of matrices are related to the second order condition for the unconstrained problems. We have also encountered the definiteness of matrices for the properties of the Slutsky matrix.

Definition 1.A.1 (Negative Definite)**.** A (symmetric) $N \times N$ matrix M is negative definite if

$$y^T M y < 0 \tag{1}$$

for all non-zero $y \in \mathbb{R}^N$.

Definition 1.A.2 (Negative Semi-definite)**.** A (symmetric) $N \times N$ matrix M is *negative semi-definite* if

$$y^T M y \leq 0 \tag{2}$$

for all $y \in \mathbb{R}^N$.

Example 1. $M = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$ is negative definite since for any non-zero $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, we have

$$\begin{aligned} y^T M y &= \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -2y_1 + y_2 & y_1 - 2y_2 + y_3 & y_2 - 2y_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ &= -\left[y_1^2 + (y_1 - y_2)^2 + (y_2 - y_3)^2 + y_3^2 \right] < 0. \end{aligned}$$

This result is the negative of sum of squares, and therefore non-positive. Furthermore, the result is zero only if $y_1 = y_2 = y_3 = 0$ that is, when y is the zero vector. Therefore, for any non-zero vector y, the result is always negative.

Example 2. $M = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ is negative semi-definite since for any $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, we have

$$y^T My = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -y_1 + y_2 & y_1 - y_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -(y_1 + y_2)^2 \leq 0.$$

This result is the negative of sum of squares, and therefore non-positive. When $y_1 = -y_2$, for example $y = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, the result is 0.

Note that a matrix M with all negative entries may not be negative definite. Example 3 illustrates the case where all entries in M is negative whereas M is not negative definite.

Example 3. $M = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$ is not negative definite since for $y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ we have

$$y^T My = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 2 > 0.$$

Similarly, we could define positive (semi-)definite matrices analogously.

Definition 1.A.3 (Positive Definite)**.** A (symmetric) $N \times N$ matrix M is positive definite if

$$y^T My > 0 \tag{3}$$

for all non-zero $y \in \mathbb{R}^N$.

Definition 1.A.4 (Positive Semi-definite)**.** A (symmetric) $N \times N$ matrix M is *positive semi-definite* if

$$y^T My \geq 0 \tag{4}$$

for all $y \in \mathbb{R}^N$.

Remark. A matrix that is not positive semi-definite and not negative semi-definite is called **indefinite**.

There are various ways to check the definiteness of matrices. In Examples 1, 2 and 3, we have used the definition to check the definiteness. Below, we will introduce the

determinantal test for definiteness. Before discussing the general theorem, we need to learn some new concepts.

Definition 1.A.5 (Principal Submatrix and Principal Minor)**.** Let M be a $N \times N$ matrix. A $k \times k$ submatrix of M formed by deleting $n-k$ rows and the same $n-k$ columns of M is called the k^{th} order **principal submatrix** of M. The determinant of a principal submatrix is called the k^{th} order **principal minor** of M.

Example 4. For a general 3×3 matrix $M = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

1. There is one 3^{rd} order principal minor, namely, $\det M$;

2. There are three 2^{nd} order principal minors, namely,

   a) $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, formed by deleting the 3^{rd} row and the 3^{rd} column;

   b) $\det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}$, formed by deleting the 2^{nd} row and the 2^{nd} column;

   c) $\det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$, formed by deleting the 1^{st} row and the 1^{st} column.

3. There are three 1^{st} order principal minors, namely,

   a) $\det \begin{bmatrix} a_{11} \end{bmatrix}$, formed by deleting the 2^{nd} and 3^{rd} rows and colomns;

   b) $\det \begin{bmatrix} a_{22} \end{bmatrix}$, formed by deleting the 1^{st} and 3^{rd} rows and colomns;

   c) $\det \begin{bmatrix} a_{33} \end{bmatrix}$, formed by deleting the 1^{st} and 2^{nd} rows and colomns.

Definition 1.A.6 (Leading Principal Submatrix and Leading Principal Minor)**.** Let M be a $N \times N$ matrix. The k^{th} order prinipal submatrix of M obtained by deleting the last $n-k$ rows and column of M is called the k^{th} order **leading principal submatrix** of M; and its determinant is called the k^{th} order **leading principal minor** of M.

Example 5. For the general 3×3 matrix in Example 4,

1. The 3^{rd} order leading principal minor is $\det M$;
2. The 2^{nd} order leading principal minor is $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$;
3. The 1^{st} order leading principal minor is $\det \begin{bmatrix} a_{11} \end{bmatrix}$.

The following two theorems provide the algorithm for testing the definiteness of a symmetric matrix.

Theorem 1.A.1. *Let M be an $N \times N$ symmetric matrix. Then*

1. *M is positive definite if and only if all its leading principal minors are positive;*
2. *M is negative definite if and only if all its leading principal minors of odd order are negative; and all its leading principal minors of even order are positive.*

Theorem 1.A.2. *Let M be an $N \times N$ symmetric matrix. Then*

1. *M is positive semi-definite if and only if all its principal minors are non-negative;*
2. *M is negative semi-definite if and only if all its principal minors of odd order are non-positive ; and all all its principal minors of even order are non-negative.*

Remark. Please note that to check the semi-definiteness of matrices, we must unfortunately check not only the leading principal minors, but all principal minors.

Remark. The second-order partial derivative matrix, F_{xx}, is called *Hessian Matrix.*

1.B. Bordered Hessian matrix

As is mentioned in class, we could use *bordered Hessian matrix* to check the second-order condition.

Definition 1.B.1. The matrix

$$\begin{bmatrix} 0 & -G_x \\ -{G_x}^T & F_{xx} - \lambda G_{xx} \end{bmatrix}$$

is called *Bordered Hessian Matrix.*

To check the second-order sufficient condition, we need to look at $n-m$ of the bordered Hessian's leading principal minors. Intuitively, we can think of the m constraints as reducing the problem to one with $n-m$ free variables.[1] The smallest minor we consider consisting of the truncated first $2m+1$ rows and columns, the next consisting of the truncated first $2m+2$ rows and columns, and so on, with the last being the determinant of the entire bordered Hessian. A sufficient condition for a local maximum of F is that the smallest minor has the same sign as $(-1)^{m+1}$ and that the rest of the principal minors alternate in sign. The result is summarized in Theorem 1.B.1 below.

Theorem 1.B.1 (Second-order Sufficient Condition for Constrained Maximization Problem)**.** *If the last $n-m$ leading principal minors of the bordered Hessian matrix at the proposed optimum x^* is such that the smallest minor (the $(2m+1)^{th}$ minor) has the same sign as $(-1)^{m+1}$ and the rest of the principal minors alternate in sign, then x^* is the local maximum of the constrained maximization problem.*

Example 6. Consider the following maximization problem with three variables ($n=3$) and two constraints ($m=2$):

$$\max_{x,y,z} F(x,y,z) \equiv z$$

$$\text{s.t. } G^1(x,y,z) \equiv x+y+z = 12$$

$$G^2(x,y,z) \equiv x^2+y^2-z = 0$$

The Lagrangian is $\mathcal{L}(x,y,z,\lambda,\mu) = z + \lambda(12-x-y-z) + \mu(-x^2-y^2+z)$.

The first-order necessary conditions are

$$\partial\mathcal{L}/\partial x = -\lambda - 2\mu x = 0$$

$$\partial\mathcal{L}/\partial y = -\lambda - 2\mu y = 0$$

$$\partial\mathcal{L}/\partial z = 1 - \lambda + \mu = 0$$

$$\partial\mathcal{L}/\partial \lambda = 12 - x - y - z = 0$$

$$\partial\mathcal{L}/\partial \mu = -x^2 - y^2 + z = 0$$

The stationary points are $(x^*, y^*, z^*, \lambda, \mu) = (2, 2, 8, \frac{4}{5}, -\frac{1}{5})$ and $(-3, -3, 18, \frac{6}{5}, \frac{1}{5})$.

[1] For example, the maximization problem: $\max_{x,y,z} x + y^2 + z$ subject to $x+y+z=1$ can be reduced to $\max_{x,y} x + y^2 + (1-x-y)$ with no constraint.

The bordered Hessian matrix is

$$\begin{bmatrix} 0 & 0 & -G_x^1 & -G_y^1 & -G_z^1 \\ 0 & 0 & -G_x^2 & -G_y^2 & -G_z^2 \\ -G_x^1 & -G_x^2 & \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} \\ -G_y^1 & -G_y^2 & \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} \\ -G_z^1 & -G_z^2 & \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2x & -2y & 1 \\ -1 & -2x & -2\mu & 0 & 0 \\ -1 & -2y & 0 & -2\mu & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

We need to check $n - m = 1$ leading principal minors, i.e., we only need to check the determinant of the bordered Hessian. For local maximum, the sign requirement is $(-1)^{m+1} = (-1)^3 < 0$.

1. For the first proposed optimum $(x^*, y^*, z^*, \lambda, \mu) = (2, 2, 8, \frac{4}{5}, -\frac{1}{5})$, the determinant of the bordered Hessian is 20;
2. For the second proposed optimum $(x^*, y^*, z^*, \lambda, \mu) = (-3, -3, 18, \frac{6}{5}, \frac{1}{5})$, the determinant of the bordered Hessian is -20.

Thus, the 2^{nd} proposed optimum $(x^*, y^*, z^*, \lambda, \mu) = (-3, -3, 18, \frac{6}{5}, \frac{1}{5})$ is a local maximum.

Example 7. Consider the following maximization problem with three variables $(n = 3)$ and one constraint $(m = 1)$:

$$\max_{x,y,z} F(x, y, z) \equiv x + y + z$$

$$\text{s.t. } G^1(x, y, z) \equiv x^2 + y^2 + z^2 = 3$$

The Lagrangian is $\mathcal{L}(x, y, z, \lambda) = x + y + z + \lambda(3 - x^2 - y^2 - z^2)$.

The first-order necessary conditions are

$$\partial\mathcal{L}/\partial x = 1 - 2\lambda x = 0$$

$$\partial\mathcal{L}/\partial y = 1 - 2\lambda y = 0$$

$$\partial\mathcal{L}/\partial z = 1 - 2\lambda z = 0$$

$$\partial\mathcal{L}/\partial \lambda = 3 - x^2 - y^2 - z^2 = 0$$

The stationary points are $(x^*, y^*, z^*, \lambda) = (-1, -1, -1, -\frac{1}{2})$ and $(1, 1, 1, \frac{1}{2})$.

The bordered Hessian matrix is

$$\begin{bmatrix} 0 & -G_x^1 & -G_y^1 & -G_z^1 \\ -G_x^1 & \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} \\ -G_y^1 & \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} \\ -G_z^1 & \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} \end{bmatrix} = \begin{bmatrix} 0 & -2x & -2y & -2z \\ -2x & -2\lambda & 0 & 0 \\ -2y & 0 & -2\lambda & 0 \\ -2z & 0 & 0 & -2\lambda \end{bmatrix}$$

We need to check $n - m = 2$ leading principal minors, i.e., the 3^{rd} order and the entire bordered Hessian. For local maximum, the sign requirement is $(-1)^{m+1} = (-1)^2 > 0$ for the 3^{rd} order leading principal minor and < 0 for the entired bordered Hessian.

1. For the first proposed optimum $(x^*, y^*, z^*, \lambda) = (-1, -1, -1, -\frac{1}{2})$, the 3^{rd} order leading principal minor is $-8 < 0$ and the determinant of the bordered Hessian is $-12 < 0$;
2. For the second proposed optimum $(x^*, y^*, z^*, \lambda) = (1, 1, 1, \frac{1}{2})$, the 3^{rd} order leading principal minor is $8 > 0$ and the determinant of the bordered Hessian is $-12 < 0$.

Thus, the 2^{nd} proposed optimum $(x^*, y^*, z^*, \lambda) = (1, 1, 1, \frac{1}{2})$ is a local maximum.