Section 5
Method of Lagrange Multipliers

Christopher Ting

Hiroshima University

✉: cting@hiroshima-u.ac.jp
☺: http://cting.x10host.com/
☎: +81 082-424-6451
: A1棟 131-1

Table of Contents

Setup

- Suppose the functions f and $g_1, g_2, \ldots, g_m$ are **continuously differentiable** on an **open set** D in $\mathbb{R}^n$.
- Suppose that $m < n$ and

$$g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0 \tag{1}$$

  on a nonempty **subset** D_1 of D, where $\mathbf{X} \in D_1$.
- If $\mathbf{X}_0 \in D_1$, and if there is a **neighborhood** N of $\mathbf{X}_0$ such that

$$f(\mathbf{X}) \leq f(\mathbf{X}_0) \tag{2}$$

  for every $\mathbf{X}$ in $N \cap D_1$, then $\mathbf{X}_0$ is a **local maximum point of f subject to the constraints** (1).

Extreme Points under Constraints

- Conversely, if (2) is replaced by

$$f(\mathbf{X}) \geq f(\mathbf{X}_0), \tag{3}$$

  then we say a **local minimum point of f subject to the constraints** (1).

- It is also called a **local extreme point** of f subject to (1).

- More briefly, we also speak of constrained local maximum, minimum, or extreme points.

- If (2) or (3) holds for all $\mathbf{X}$ in D_1, we omit “local.”

Critical Point

- We say that $\mathbf{X}_0 = (x_{10}, x_{20}, \ldots, x_{n0})$ is a **critical point** of a **differentiable function** $L = L(x_1, x_2, \ldots, x_n)$ if

$$L_{x_i}(x_{10}, x_{20}, \ldots, x_{n0}) = 0, \quad 1 \leq i \leq n.$$

- Therefore, every **local extreme point** of L is a critical point of L.

- On the other hand, a **critical point** of L is not necessarily a local extreme point of L.

Constrained Solution

- Suppose that the system (1) of **simultaneous equations** can be solved for $x_1, \ldots, x_m$ in terms of the $x_{m+1}, \ldots, x_n$; thus

$$x_j = h_j(x_{m+1}, \ldots, x_n), \quad 1 \leq j \leq m. \tag{4}$$

- Then a **constrained extreme value** of f is an **unconstrained extreme value** of

$$f(h_1(x_{m+1}, \ldots, x_n), \ldots, h_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n). \tag{5}$$

- However, it may be difficult or impossible to find explicit formulas for $h_1, h_2, \ldots, h_m$.
- Even if it is possible, the **composite function** (5) is almost always complicated.
- Is there a better way to to find **constrained extrema**?

Theorem 1

Theorem 1: Lagrange Multipliers

Suppose that $n > m$. Suppose $\mathbf{X}_0$ is a **local extreme point** of f subject to $g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0$ and the **determinant**

$$\begin{vmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_{r_m}} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_m}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_1}} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_2}} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_{r_m}} \end{vmatrix} \neq 0 \tag{6}$$

for at least one choice of $r_1 < r_2 < \cdots < r_m$.

Theorem 1 (cont'd)

Theorem 1: Lagrange Multipliers

- Then there are constants $\lambda_1, \lambda_2, \ldots, \lambda_m$ called the **Lagrange multipliers** such that $\mathbf{X}_0$ is a **critical point** of

$$f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m.$$

- That is,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} - \lambda_1 \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} - \lambda_2 \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} - \cdots - \lambda_m \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} = 0,$$

for $1 \le i \le n$.

Method of Lagrange Multipliers

(a) Find the **critical points** of

$$f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m,$$

treating $\lambda_1, \lambda_2, \ldots, \lambda_m$ as unspecified constants.

(b) Find $\lambda_1, \lambda_2, \ldots, \lambda_m$ so that the critical points obtained in (a) satisfy the constraints.

(c) Determine which of the critical points are **constrained extreme points** of f.

(d) If a and $b_1, b_2, \ldots, b_m$ are nonzero constants and c is an arbitrary constant, then the **local extreme points** of f subject to $g_1 = g_2 = \cdots = g_m = 0$ are the same as the local extreme points of $af - c$ subject to

$$b_1 g_1 = b_2 g_2 = \cdots = b_m g_m = 0.$$

(e) Therefore, to simplify computations, we can replace $f - \lambda_1 g_1 - \lambda_2 g_2 - \cdots - \lambda_m g_m$ by the **Lagrangian**

$$L := af - \lambda_1 b_1 g_1 - \lambda_2 b_2 g_2 - \cdots - \lambda_m b_m g_m - c.$$

Theorem 2

Theorem 2: A Special Case for $m = 1$ Constraint

- Suppose that $n > 1$ and that $\mathbf{X}_0$ is a **local extreme point** of f subject to $g(\mathbf{X}) = 0$.
- We also assume that the **partial derivative** $g_{x_r}(\mathbf{X}_0) \neq 0$ for some $r \in \{1, 2, \ldots, n\}$.
- Then there is a constant λ such that

$$f_{x_i}(\mathbf{X}_0) - \lambda g_{x_i}(\mathbf{X}_0) = 0, \quad \text{for } 1 \leq i \leq n. \tag{7}$$

- In other words, $\mathbf{X}_0$ is a **critical point** of the **Lagrangian** $f - \lambda g$.

Proof of Theorem 2

- Without loss of generality, let $r = 1$. Denote

$$\mathbf{U} = (x_2, x_3, \ldots x_n) \text{ and } \mathbf{U}_0 = (x_{20}, x_{30}, \ldots x_{n0}).$$

- Since $g_{x_1}(\mathbf{X}_0) \neq 0$, here is a unique **continuously differentiable function** $h = h(\mathbf{U})$ defined on a **neighborhood** $N \subset \mathbb{R}^{n-1}$ of $\mathbf{U}_0$, such that $\big(h(\mathbf{U}), \mathbf{U}\big) \in D$ for all $\mathbf{U} \in N$, $h(\mathbf{U}_0) = x_{10}$, and

$$g(h(\mathbf{U}), \mathbf{U}) = 0, \quad \mathbf{U} \in N. \tag{8}$$

- Now define

$$\lambda = \frac{f_{x_1}(\mathbf{X}_0)}{g_{x_1}(\mathbf{X}_0)}, \tag{9}$$

which is permissible, since $g_{x_1}(\mathbf{X}_0) \neq 0$.

Proof of Theorem 2 (cont'd)

- Differentiating (8) with respect to x_i yields, for $i = 2, 3, \ldots n$,

$$\frac{\partial g(h(\mathbf{U}), \mathbf{U})}{\partial x_i} + \frac{\partial g(h(\mathbf{U}), \mathbf{U})}{\partial x_1} \frac{\partial h(\mathbf{U})}{\partial x_i} = 0, \quad \mathbf{U} \in N. \tag{10}$$

- Also,

$$\frac{\partial f(h(\mathbf{U}), \mathbf{U}))}{\partial x_i} = \frac{\partial f(h(\mathbf{U}), \mathbf{U})}{\partial x_i} + \frac{\partial f(h(\mathbf{U}), \mathbf{U})}{\partial x_1} \frac{\partial h(\mathbf{U})}{\partial x_i}, \quad \mathbf{U} \in N. \tag{11}$$

- Since $\big(h(\mathbf{U}_0), \mathbf{U}_0\big) = \mathbf{X}_0$, (10) implies that

$$\frac{\partial g(\mathbf{X}_0)}{\partial x_i} + \frac{\partial g(\mathbf{X}_0)}{\partial x_1} \frac{\partial h(\mathbf{U}_0)}{\partial x_i} = 0. \tag{12}$$

Proof of Theorem 2 (cont'd)

- If $\mathbf{X}_0$ is a **local extreme point** of f subject to $g(\mathbf{X}) = 0$, then $\mathbf{U}_0$ is an **unconstrained local extreme point** of $f\big(h(\mathbf{U}), \mathbf{U}\big)$.
- Therefore, (11) implies that

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} + \frac{\partial f(\mathbf{X}_0)}{\partial x_1}\frac{\partial h(\mathbf{U}_0)}{\partial x_i} = 0. \tag{13}$$

- Recall that a **linear homogeneous system**

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{14}$$

has a **nontrivial solution** if and only if $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$.

Proof of Theorem 2 (cont'd)

- (12) and (13) imply that

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} \\[2ex] \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{vmatrix} = 0, \quad \text{so} \quad \begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{vmatrix} = 0,$$

  since the **determinant** of a matrix and that of its **transpose** are equal.

- Therefore, the system $\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ has a **nontrivial solution**.

Proof of Theorem 2 (cont'd)

- Since $g_{x_1}(\mathbf{X}_0) \neq 0$, u must be nonzero in a **nontrivial solution**.
- Hence, we may assume that $u = 1$, so

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} 1 \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{15}$$

- In particular

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_1} + v\frac{\partial g(\mathbf{X}_0)}{\partial x_1} = 0, \quad \text{so} \quad -v = \frac{f_{x_1}(\mathbf{X}_0)}{g_{x_1}(\mathbf{X}_0)}.$$

Proof of Theorem 2 (cont'd)

- Now (9) implies that $-v = \lambda$, and (15) becomes

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_i} \\[2ex] \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g(\mathbf{X}_0)}{\partial x_1} \end{bmatrix} \begin{bmatrix} 1 \\ -\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

- Computing the topmost entry of the vector on the left yields (7).

□

Example 1

- Find the **extreme values** of $\sum_{i=1}^{n} x_i$ subject to $\sum_{i=1}^{n} x_i^2 = 1$.
- Solution: Let $L = \sum_{i=1}^{n} x_i - \frac{\lambda}{2} \sum_{i=1}^{n} x_i^2$.
- Then $L_{x_i} = 1 - \lambda x_i$, so $x_{i0} = \frac{1}{\lambda}$, $\quad 1 \le i \le n$.
- Hence $\sum_{i=1}^{n} x_{i0}^2 = n/\lambda^2$, implying $\lambda = \pm\sqrt{n}$ and

$$(x_{10}, x_{20}, \ldots, x_{n0}) = \pm \left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}} \right).$$

- Therefore, the **constrained maximum** is $\sqrt{n}$ and the **constrained minimum** is $-\sqrt{n}$.

Example 2

- Show that if

$$\frac{1}{p}+\frac{1}{q}=1, p>0, \text{ and } q>0, \tag{16}$$

then $x^{1/p}y^{1/q} \le \dfrac{x}{p}+\dfrac{y}{q}, \quad x, y \ge 0.$

- We first find the maximum of

$$f(x,y)=x^{1/p}y^{1/q}$$

subject to

$$\frac{x}{p}+\frac{y}{q}=\sigma, \quad x \ge 0, \quad y \ge 0, \tag{17}$$

where σ is a fixed but arbitrary positive number.

Example 2 (cont'd)

- Define the **Lagrangian** $L = x^{1/p}y^{1/q} - \lambda\left(\frac{x}{p} + \frac{y}{q}\right)$.
- Then, $L_x = \frac{1}{px}f(x,y) - \frac{\lambda}{p}$ and $L_y = \frac{1}{qy}f(x,y) - \frac{\lambda}{q}$.
- Setting both derivatives to zero, we find $x_0 = y_0 = f(x_0, y_0)/\lambda$.
- (16) and (17) suggest that $x_0 = y_0 = \sigma$.
- It follows that

$$f(x,y) \leq f(\sigma,\sigma) = \sigma^{1/p}\sigma^{1/q} = \sigma = \frac{x}{p} + \frac{y}{q}.$$

□

Eigenvalue and Eigenvector

- We write $\mathbf{x}' = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$.
- An **eigenvalue** of a $n \times n$ **square matrix** $\mathbf{A}$ is a number λ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, equivalently

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.$$

- This **characteristic equation** has a solution $\mathbf{x} \neq \mathbf{0}$ known as the **eigenvector**.
- From linear algebra, λ is an **eigenvalue** of $\mathbf{A}$ if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

- Suppose that $\mathbf{A}$ is **symmetric**, i.e., $(a_{ij} = a_{ji}, 1 \leq i, j \leq n)$. In this case

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n),$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real numbers.

Quadratic Form

- The function $Q(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ is a **quadratic form**.
- To find its maximum or minimum subject to $\sum_{i=1}^{n} x_i^2 = 1$, we form the **Lagrangian**:

$$L = Q(\mathbf{x}) - \lambda \sum_{i=1}^{n} x_i^2.$$

- Then $L_{x_i} = 2\sum_{j=1}^{n} a_{ij} x_j - 2\lambda x_i = 0, \quad 1 \le i \le n$. Setting the equation to 0, we obtain

$$\sum_{j=1}^{n} a_{ij} x_{j0} = \lambda x_{i0}, \quad 1 \le i \le n.$$

Connection to Eigenvector

- Therefore $\mathbf{x}_0' = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}_0$ is a **constrained critical point** of the function $Q(\mathbf{x})$ subject to $\sum_{i=1}^{n} x_i^2 = 1$ if and only if $\mathbf{A}\mathbf{x}_0 = \lambda \mathbf{x}_0$ for some λ.

- If $\mathbf{A}\mathbf{x}_0 = \mathbf{x}_0$ and $\sum_{i=1}^{n} x_{i0}^2 = 1$, then

$$Q(\mathbf{x}_0) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_{j0} \right) x_{i0} = \sum_{i=1}^{n} (\lambda x_{i0}) x_{i0} = \lambda \sum_{i=1}^{n} x_{i0}^2 = \lambda.$$

- Therefore, the largest and smallest **eigenvalues** of $\mathbf{A}$ are the maximum and minimum values of Q subject to $\sum_{i=1}^{n} x_i^2 = 1$.

Step 1

- For convenience, let $1 \leq \ell \leq m$ and suppose the **determinant** (6) is non-zerro.

- Denote $\mathbf{U} = (x_{m+1}, x_{m+2}, \ldots x_n)$ and $\mathbf{U}_0 = (x_{m+1,0}, x_{m+2,0}, \ldots x_{n0})$.

- In view of (6), there are unique **continuously differentiable functions** $h_\ell = h_\ell(\mathbf{U})$ defined on a **neighborhood** N of $\mathbf{U}$, such that

$$\big(h_1(\mathbf{U}), h_2(\mathbf{U}), \ldots, h_m(\mathbf{U}), \mathbf{U}\big) \in D, \quad \text{for all} \quad \mathbf{U} \in N,$$

$$(h_1(\mathbf{U_0}), h_2(\mathbf{U_0}), \ldots, h_m(\mathbf{U_0}), \mathbf{U}_0) = \mathbf{X}_0, \tag{18}$$

and

$$g_\ell\big(h_1(\mathbf{U}), h_2(\mathbf{U}), \ldots, h_m(\mathbf{U}), \mathbf{U}\big) = 0, \quad \mathbf{U} \in N, \quad 1 \leq \ell \leq m. \tag{19}$$

Step 2

- Again from (6), we can consider the **system of equations**

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \\ \vdots \\ f_{x_m}(\mathbf{X}_0) \end{bmatrix}. \tag{20}$$

- It has a **unique solution**, implying that, for $1 \le i \le m$,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} - \lambda_1 \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} - \lambda_2 \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} - \cdots - \lambda_m \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} = 0. \tag{21}$$

Step 3

- For $m+1 \leq i \leq n$, differentiating (19) with respect to x_i and recalling (18) yields,

$$\frac{\partial g_\ell(\mathbf{X}_0)}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial g_\ell(\mathbf{X}_0)}{\partial x_j} \frac{\partial h_j(\mathbf{X}_0)}{\partial x_i} = 0, \quad 1 \leq \ell \leq m.$$

- If $\mathbf{X}_0$ is a **local extreme point** f subject to $g_1(\mathbf{X}) = g_2(\mathbf{X}) = \cdots = g_m(\mathbf{X}) = 0$, then $\mathbf{U}_0$ is an **unconstrained local extreme point** of $f\big(h_1(\mathbf{U}), h_2(\mathbf{U}), \ldots h_m(\mathbf{U}), \mathbf{U}\big)$.

- Therefore,

$$\frac{\partial f(\mathbf{X}_0)}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f(\mathbf{X}_0)}{\partial x_j} \frac{\partial h_j(\mathbf{X}_0)}{\partial x_i} = 0.$$

Step 4

- The last two equations imply that

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{vmatrix} = 0.$$

Step 5

- So the **determinant** of the **transposed matrix** is also

$$\begin{vmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{vmatrix} = 0.$$

Step 6

- Therefore, there are constant $c_0, c_1, \ldots, c_m$, not all zero, such that

$$\begin{bmatrix} \frac{\partial f(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_3 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{22}$$

Step 7

- If $c_0 = 0$, then

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- Since the dxdeterminant is non-zero, it must be that $c_1 = c_2 = \cdots = c_m = 0$.

Step 8

- Hence, we may assume that $c_0 = 1$ in a nontrivial solution of (22). Therefore,

$$\begin{bmatrix} \dfrac{\partial f(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{23}$$

Step 9

- It implies that

$$\begin{bmatrix} \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_1(\mathbf{X}_0)}{\partial x_m} \\ \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_2(\mathbf{X}_0)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_1} & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_2} & \cdots & \dfrac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} -c_1 \\ -c_2 \\ \vdots \\ -c_m \end{bmatrix} = \begin{bmatrix} f_{x_1}(\mathbf{X}_0) \\ f_{x_2}(\mathbf{X}_0) \\ \vdots \\ f_{x_m}(\mathbf{X}_0) \end{bmatrix}.$$

Step 10

- Since (22) has only one solution, implying that $c_j = -\lambda, 1 \leq j \leq n$. So (23) becomes

$$\begin{bmatrix} \frac{\partial f(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_i} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_i} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_i} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_1} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_1} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_1} \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_2} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_1(\mathbf{X}_0)}{\partial x_m} & \frac{\partial g_2(\mathbf{X}_0)}{\partial x_m} & \cdots & \frac{\partial g_m(\mathbf{X}_0)}{\partial x_m} \end{bmatrix} \begin{bmatrix} 1 \\ -\lambda_1 \\ -\lambda_2 \\ \vdots \\ -\lambda_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

- Computing the topmost entry of the vector on the left yields (21), which completes the proof.

Keywords