SECOND DERIVATIVE TEST FOR CONSTRAINED EXTREMA

This handout presents the second derivative test for a local extrema of a Lagrange multiplier problem. The Section 1 presents a geometric motivation for the criterion involving the second derivatives of both the function f and the constraint function g. The main result is given in section 3, with the special cases of one constraint given in Sections 4 and 5 for two and three dimensions respectively. The result is given in terms of the determinant of what is called the bordered Hessian matrix, which is defined in Section 2 using the Lagrangian function.

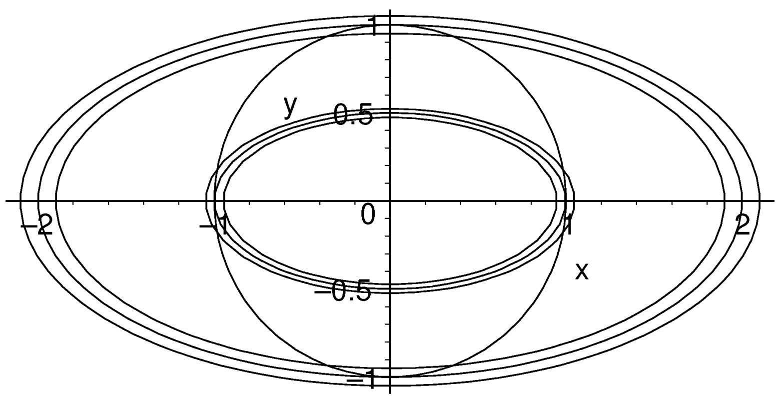
1. **Intuitive Reason for Terms in the Test**

In order to understand why the conditions for a constrained extrema involve the second partial derivatives of both the function maximized f and the constraint function g, we start with an example in two dimensions.

Consider the extrema of $f(x, y) = x^2 + 4y^2$ on the constraint $1 = x^2 + y^2 = g(x, y)$. The four critical points found by Lagrange multipliers are $(\pm 1, 0)$ and $(0, \pm 1)$. The points $(\pm 1, 0)$ are minima, $f(\pm 1, 0) = 1$; the points $(0, \pm 1)$ are maxima, $f(0, \pm 1) = 2$. The Hessian of f is the same for all points,

$$Hf(x, y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}.$$

Therefore the fact that some of the critical points are local minima and others are local maxima cannot depend on the second partial derivatives of f alone.



The above figure displays the level curves $1 = g(x, y)$, and $f(x, y) = C$ for $C = (0.9)^2$, 1, $(1.1)^2$, (1.9), 2^2, and $(2.1)^2$ on one plot.

The level curve $f(x, y) = 1$ intersects $1 = g(x, y)$ at $(\pm 1, 0)$. For nearby values of f, the level curve $f(x, y) = (0.9)^2$ does not intersect $1 = g(x, y)$, while the level curve $f(x, y) = (1.1)^2$ intersects $1 = g(x, y)$ in four points. Therefore, the points $(\pm 1, 0)$ are local minima for f. Notice that the level curve $f(x, y) = 1$ bends more sharply near $(\pm 1, 0)$ than the level curve for g and so the level curve for f lies inside the level curve for g. Since it lies inside the level curve for g and the gradient of f points outward, these points are local minima for f on the level curve of g.

On the other hand, the level curve $f(x, y) = 4$ intersects $1 = g(x, y)$ at $(0, \pm 1)$, $f(x, y) = (1.9)$ intersects in four points, and $f(x, y) = (2.1)^2$ does not intersect. Therefore, the points $(0, \pm 1)$ are local maxima for f. Notice that the level curve $f(x, y) = 4$ bends less sharply near $(0, \pm 1)$ than the level curve for g and so the level curve for f lies outside the level curve for g. Since it lies outside the level curve for g and the gradient of f points outward, these points are local maxima for f on the level curve of g.

Thus, the second partial derivatives of f are the same at both $(\pm 1, 0)$ and $(0, \pm 1)$, but the sharpness with which the two level curves bend determines which are local maxima and which are local minima. This discussion motivates the fact that it is the comparison of the second partial derivatives of f and g which is relevant.

2. **Lagrangian Function**

One way to getting the relevant matrix is to form the Lagrangian function, which is a combination of f and g. For the problem of finding the extrema (maxima or minima) of $f(\mathbf{x})$ with ik constraints $g_\ell(\mathbf{x}) = C_\ell$ for $1 \leq \ell \leq k$, the *Lagrangian function* is defined to be the function

$$L(\lambda, \mathbf{x}) = f(\mathbf{x}) - \sum\nolimits_{\ell=1}^{k} \lambda_\ell \left[g_\ell(\mathbf{x}) - C_\ell \right].$$

The solution of the Lagrange multiplier problems is then a critical point of L,

$$\frac{\partial L}{\partial \lambda_\ell}(\lambda^*, \mathbf{x}^*) = -g_\ell(\mathbf{x}^*) + C_\ell = 0, \qquad \text{for } 1 \leq \ell \leq k \text{ and}$$

$$\frac{\partial L}{\partial x_i}(\lambda^*, \mathbf{x}^*) = \frac{\partial f}{\partial x_i}(\mathbf{x}^*) - \sum\nolimits_{\ell=1}^{k} \lambda_\ell^* \frac{\partial g_\ell}{\partial x_i}(\mathbf{x}^*) = 0 \qquad \text{for } 1 \leq i \leq n.$$

The second derivative test involves the matrix of all second partial derivatives of L, including those with respect to λ. In dimensions n greater than two, the test also involves submatrices. Notice that $\frac{\partial^2 L}{\partial \lambda^2}(\lambda^*, \mathbf{x}^*) = 0$ and $\frac{\partial^2 L}{\partial \lambda \ell \partial x_i}(\lambda^*, \mathbf{x}^*) = -\frac{\partial g_\ell}{\partial x_i}(\mathbf{x}^*)$. We could use $\frac{\partial g_\ell}{\partial x_i}(\mathbf{x}^*)$ in the matrix instead of $-\frac{\partial g_\ell}{\partial x_i}(\mathbf{x}^*)$ it does not change the determinant (both a row and a column are multiplied by minus one). The matrix of all second partial derivatives of L is called the bordered Hessian matrix because the the second derivatives of L with respect to the x_i variables is bordered by the first order partial derivatives of g. The *bordered Hessian matrix* is defined to be

$$HL(\lambda^*, \mathbf{x}^*) = \begin{bmatrix} 0 & \cdots & 0 & -(g_1)_{x_1} & \dots & -(g_1)_{x_n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -(g_k)_{x_1} & \dots & -(g_k)_{x_n} \\ -(g_1)_{x_1} & \cdots & -(g_k)_{x_1} & L_{x_1 x_1} & \dots & L_{x_1 x_n} \\ \vdots & \vdots & & \vdots & & \\ -(g_1)_{x_n} & \cdots & -(g_k)_{x_n} & L_{x_n x_1} & \dots & L_{x_n x_n} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -Dg \\ -Dg^T & D_{\mathbf{x}} L \end{bmatrix} \tag{1}$$

where all the partial derivatives are evaluated with $\mathbf{x} = \mathbf{x}^*$ and $\lambda = \lambda^*$. In the following, we use the notation $D^2_{\mathbf{x}} L^* = D^2_{\mathbf{x}} L(\lambda^*, \mathbf{x}^*) = D^2 f(\mathbf{x}^*) - \sum_{\ell=1}^{k} \lambda_\ell^* D^2(g_\ell)(\mathbf{x}^*)$ for this submatrix that appears in the bordered Hessian.

3. **Derive Second Derivative Conditions**

The first section gave an intuitive reason why the second derivative test should involve the second derivatives of the constraint as well as the function being extremized. In this section, we derive the exact condition which involves the bordered Hessian defined in the last section. First, we should what the second derivative of f along a curve in the level of of the constraint function g. Then, we apply the result mentioned for a quadratic form on the null space of a linear map.

Lemma 1. *Assume that $\mathbf{x}^*$ and $\lambda^* = (\lambda_1^*, \dots, \lambda_k^*)$ meet the first-order conditions of an extrema of f on the level set $g_\ell(\mathbf{x}) = C_\ell$ for $1 \leq \ell \leq k$. If $\mathbf{r}(t)$ is a curve in $g^{-1}(\mathbf{C})$ with $\mathbf{r}(0) = \mathbf{x}^*$ and $\mathbf{r}'(0) = \mathbf{v}$, then*

$$\left. \frac{d^2}{dt^2} f(\mathbf{r}(t)) \right|_{t=0} = \mathbf{v}^T \left[D^2 f(\mathbf{x}^*) - \sum_{\ell=1}^{k} \lambda_\ell^* D^2 g_\ell(\mathbf{x}^*) \right] \mathbf{v} = \mathbf{v}^T D^2_{\mathbf{x}} L(\lambda^*, \mathbf{x}^*) \mathbf{v}.$$

Proof. Using the chain rule and product rule,

$$\frac{d}{dt} f(\mathbf{r}(t)) = Df(\mathbf{r}(t))\mathbf{r}'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{r}(t)) r_i'(t)$$
(by the chain rule)

and

$$\begin{aligned} \frac{d^2}{dt^2} f(\mathbf{r}(t))\bigg|_{t=0} &= \sum_{i=1}^{n} \frac{d}{dt} \frac{\partial f}{\partial x_i}(\mathbf{r}(t))\bigg|_{t=0} r_i'(0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{r}(0)) r_i''(0) \\ &\qquad \text{(by the product rule)} \\ &= \sum_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^*) r_i'(0) r_j'(0) + Df(\mathbf{x}^*)\mathbf{r}''(0) \\ &\qquad \text{(by the chain rule)} \\ &= (\mathbf{r}'(0))^T D^2 f(\mathbf{x}^*)\mathbf{r}'(0) + \sum_{\ell=1}^{k} \lambda_\ell^* D(g_\ell)(\mathbf{x}^*)\mathbf{r}''(0). \end{aligned}$$

In the last equality, we used the definition of $D^2 f$ and the fact that $Df(\mathbf{x}^*) = \sum_{\ell=1}^{k} \lambda_\ell^* D(g_\ell)(\mathbf{x}^*)$.

We can perform a similar calculation for the constraint equation $0 = g_\ell(\mathbf{r}(t))$ whose derivatives are zero:

$$\begin{aligned} 0 &= \frac{d}{dt} g_\ell(\mathbf{r}(t)) = \sum_{i=1,\dots,n} \left(\frac{\partial g_\ell}{\partial x_i}(\mathbf{r}(t))\right) \mathbf{r}_i'(t), \\ 0 &= \frac{d^2}{dt^2} g_\ell(\mathbf{r}(t))\bigg|_{t=0} = \sum_{i=1,\dots,n} \frac{d}{dt} \left(\frac{\partial g_\ell}{\partial x_i}(\mathbf{r}(t))\right) \mathbf{r}_i'(t)\bigg|_{t=0} \\ &= \sum_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \left(\frac{\partial^2 g_\ell}{\partial x_j \partial x_i}(\mathbf{x}^*)\right) r_i'(0) r_j'(0) + D(g_\ell)(\mathbf{x}^*)\mathbf{r}''(0), \qquad \text{and} \end{aligned}$$

$$\lambda_\ell^* D(g_\ell)(\mathbf{x}^*)\mathbf{r}''(0) = -\lambda_\ell^* (\mathbf{r}'(0))^T D^2(g_\ell)(\mathbf{x}^*)\mathbf{r}'(0).$$

Substituting this equality into the expression for the second derivative of $f(\mathbf{r}(t))$,

$$\frac{d^2}{dt^2} f(\mathbf{r}(t))\bigg|_{t=0} = \mathbf{v}^T \left[D^2 f(\mathbf{x}^*) - \sum_{\ell=1,\dots,k} \lambda_\ell^* D^2 g_\ell(\mathbf{x}^*) \right] \mathbf{v},$$

where $\mathbf{v} = \mathbf{r}'(0)$. This is what is claimed. □

The next theorem uses the above lemma to derive conditions for local maxima and minima in terms of the second derivative of the Lagrangian $D_{\mathbf{x}}^2 L^*$ on the set of vectors $\text{Nul}(Dg(\mathbf{x}^*))$.

Theorem 2. *Assume $f, g_\ell : \mathbb{R}^n \to \mathbb{R}$ are C^2 for $1 \leq \ell \leq k$. Assume that $\mathbf{x}^* \in \mathbb{R}^n$ and $\lambda^* = (\lambda_1^*. \dots, \lambda_k^*)$ meet the first-order conditions of the Theorem of Lagrange on $g^{-1}(\mathbf{C})$.*

- **a.** *If f has a local maximum on $g^{-1}(\mathbf{C})$ at $\mathbf{x}^*$, then $\mathbf{v}^T D_{\mathbf{x}}^2 L^* \mathbf{v} \leq 0$ for all $\mathbf{v} \in Nul(Dg(\mathbf{x}^*))$.*
- **b.** *If f has a local minimum on $g^{-1}(\mathbf{C})$ at $\mathbf{x}^*$, then $\mathbf{v}^T D_{\mathbf{x}}^2 L^* \mathbf{v} \geq 0$ for all $\mathbf{v} \in Nul(Dg(\mathbf{x}^*))$.*
- **c.** *If $\mathbf{v}^T D_{\mathbf{x}}^2 L^* \mathbf{v} < 0$ for all $\mathbf{v} \in Nul(Dg(\mathbf{x}^*)) \setminus \{\mathbf{0}\}$, then $\mathbf{x}^*$ is a strict local maximum of f on $g^{-1}(\mathbf{C})$.*
- **d.** *If $\mathbf{v}^T D_{\mathbf{x}}^2 L^* \mathbf{v} > 0$ for all $\mathbf{v} \in Nul(Dg(\mathbf{x}^*)) \setminus \{\mathbf{0}\}$, then $\mathbf{x}^*$ is a strict local minimum of f on $g^{-1}(\mathbf{C})$.*
- **e.** *If $\mathbf{v}^T D_{\mathbf{x}}^2 L^* \mathbf{v}$ is positive for some vector $\mathbf{v} \in Nul(Dg(\mathbf{x}^*))$ and negative for another such vector, then $\mathbf{x}^*$ is neither a local maximum nor a local minimum of f on $g^{-1}(\mathbf{C})$.*

Proof. (**b**) We consider the case of minima. (The case of maximum just reverses the direction of the inequality.) Lemma 1 shows that

$$\frac{d^2}{dt^2} f(\mathbf{r}(t))\bigg|_{t=0} = \mathbf{v}^T D^2_{\mathbf{x}} L^* \mathbf{v},$$

where $\mathbf{v} = \mathbf{r}'(0)$. If $\mathbf{x}^*$ is a local minimum on $g^{-1}(\mathbf{C})$ then

$$\frac{d^2}{dt^2} f(\mathbf{r}(t))\bigg|_{t=0} \geq 0$$

for any curves $\mathbf{r}(t)$ in $g^{-1}(\mathbf{C})$ with $\mathbf{r}(0) = \mathbf{x}^*$. Thus, $\mathbf{v}^T D^2_{\mathbf{x}} L^* \mathbf{v} \geq 0$ for any vector $\mathbf{v}$ tangent to a curve in $g^{-1}(\mathbf{C})$. But the implicit function theorem implies that these are the same as the vector in the null space $\text{Nul}(Dg(\mathbf{x}^*))$.

(**d**) If $\mathbf{v}^T D^2_{\mathbf{x}} L^* \mathbf{v} > 0$ for all vectors $\mathbf{v} \neq \mathbf{0}$ in $\text{Nul}(Dg(\mathbf{x}^*))$, then

$$\frac{d^2}{dt^2} f(\mathbf{r}(t))\bigg|_{t=0} = \mathbf{r}'(0)^T D^2_{\mathbf{x}} L^* \, \mathbf{r}'(0) > 0$$

for any curves $\mathbf{r}(t)$ in $g^{-1}(\mathbf{C})$ with $\mathbf{r}(0) = \mathbf{x}^*$ and $\mathbf{r}'(0) \neq \mathbf{0}$. This latter condition implies that $\mathbf{x}^*$ is a strict local minimum on $g^{-1}(\mathbf{C})$.

For part (**e**), if $\mathbf{v}^T D^2_{\mathbf{x}} L^* \mathbf{v}$ is both positive and negative, then there are some curves where the value of f is greater than at $\mathbf{x}^*$ and others on which the value is less. □

If we combine this result with the conditions we gave for the maximum of a quadratic form on the null space of a linear map, we get the theorem given in the book.

Combining with the earlier theorem on constrained quadratic forms, we get the following theorem given in the book.

Theorem 3. *Assume that $f, g_\ell : \mathbb{R}^n \to \mathbb{R}$ are C^2 for $1 \leq \ell \leq k$ and that $\lambda^* = (\lambda^*_1, \dots, \lambda^*_l)$ and $\mathbf{x}^*$ satisfied the first order conditions for a extrema of f on $g^{-1}(\mathbf{C})$. Assume that the $k \times k$ submatrix of $Dg(\mathbf{x}^*)$ formed by the first k columns has nonzero determinant,* $\det\left(\frac{\partial g_\ell}{\partial x_j}(\mathbf{x}^*)\right)_{1 \leq i,j \leq k} \neq 0$. *Let $\mathbf{H}_j$ be the upper left $j \times j$ submatrix of $HL(\lambda^*, \mathbf{x}^*)$.*

(**a**) *If $(-1)^k \det(H_j) > 0$ for $2k+1 \leq j \leq n+k$, the the function f has a local minimum at $\mathbf{x}^*$ on the level set $g^{-1}(\mathbf{C})$. (Notice that the sign given by $(-1)^k$ depends on the rank k and not j.)*

(**b**) *If $(-1)^{j-k} \det(H_j) > 0$ for $2k+1 \leq j \leq n+k$, the the function f has a local maximum at $\mathbf{x}^*$ on the level set $g^{-1}(\mathbf{C})$. Notice that the sign given by $(-1)^{j-k}$ depends on j and alternates sign. The condition on the signs of the determinants can be express as $(-1)^k \det(H_{2k+1}) < 0$, and the rest of the sequence $(-1)^k \det(H_j)$ alternate signs with j.*

(**c**) *If these determinants $(-1)^k \det(H_j) \neq 0$ for $2k+1 \leq j \leq n+k$ but fall into a different pattern of signs than the above two cases, then the critical point is some type of saddle.*

Remark 1. Notice that the null space $\text{Nul}(Dg(\mathbf{x}^*))$ had dimension $n-k$, so we need $n-k$ conditions. The range of j in the assumptions of the theorem contains $n-k$ values.

In the case of negative definite, the first case for $j = 2k+1$, $(-1)^k \det(H_j) < 0$ and the terms $(-1)^k \det(H_j) < 0$ alternate sign.

4. One Constraint in Two Dimensions

Now we turn to the case of two variables and one constraint, and consider a extrema of a function $f(x,y)$ on a constraint $C = g(x,y)$. Assume that (x^*, y^*) and λ^* satisfy the equations for a critical point of the Lagrangian equations

$$\nabla f_{(x^*,y^*)} = \lambda^* \nabla g_{(x^*,y^*)} \qquad \text{and} \qquad C = g(x^*, y^*). \tag{2}$$

Let

$$L(\lambda, x, y) = f(x, y) - \lambda\ [g(x, y) - C]$$

be the Lagrangian function for the problem. Form the *bordered Hessian matrix*

(3)
$$\boxed{HL = \begin{pmatrix} 0 & -g_x & -g_y \\ -g_x & L_{xx} & L_{xy} \\ -g_y & L_{yx} & L_{yy} \end{pmatrix},}$$

where the partial derivatives are evaluated at (x^*, y^*) and λ^*.

The following theorem then contains the statement of the result for local extrema.

Theorem 4. *Let f and g be real valued functions on $\mathbb{R}^2$. Let $(x^*, y^*) \in \mathbb{R}^2$ and λ^* be a solution of the Lagrange multiplier problem $\nabla f_{(x^*,y^*)} = \lambda^* \nabla g_{(x^*,y^*)}$ and $C = g(x^*, y^*)$. Define the bordered matrix HL by equation* (3)*.*

(a) *The point (x^*, y^*) is a local minimum of f on $C = g(x, y)$, if $-\det(HL(\lambda^*, x^*, y^*)) > 0$.*
(b) *The point (x^*, y^*) is a local maximum of f on $C = g(x, y)$, if $-\det(HL(\lambda^*, x^*, y^*)) < 0$.*

The minus one before the determinant comes from the fact that there is one constraint.

Example 1. Consider the example

$$f(x, y) = x^2 + 4y^2 \qquad \text{and}$$
$$g(x, y) = x^2 + y^2 = 1.$$

The equations to solve for Lagrange multipliers are

$$2x = \lambda 2x,$$
$$8y = \lambda 2y, \qquad \text{and}$$
$$1 = x^2 + y^2.$$

Solving these yields (i) $x = 0$, $y = \pm 1$, and $\lambda = 4$, and (ii) $x = \pm 1$, $y = 0$, and $\lambda = 1$.

The Lagrangian function is

$$L(\lambda, x, y) = x^2 + 4y^2 - \lambda(x^2 + y^2) + \lambda.$$

The bordered Hessian matrix is

$$HL = \begin{pmatrix} 0 & -2x & -2y \\ -2x & 2 - 2\lambda & 0 \\ -2y & 0 & 8 - 2\lambda \end{pmatrix}.$$

(i) At the first pair of points, $x = 0$, $y = \pm 1$, and $\lambda = 4$,

$$HL(4, 0, \pm 1) = \begin{pmatrix} 0 & 0 & \mp 2 \\ 0 & -6 & 0 \\ \mp 2 & 0 & 0 \end{pmatrix}.$$

So, $-\det(HL) = -(-1)(-6)(\pm 2)^2 = -24 < 0$, and these points are local maxima.

(ii) At the second pair of points $x = \pm 1$, $y = 0$, and $\lambda = 1$,

$$HL(1, \pm 1, 0) = \begin{pmatrix} 0 & \mp 2 & 0 \\ \mp 2 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

So, $-\det(HL) = -(-1)(6)(\pm 2)^2 = 24 > 0$, and these points are local minima.

These results agree with the answers found by taking the values at the points, $f(\pm 1, 0) = 1$ and $f(0, \pm 1) = 4$.

5. **One Constraint in Three Dimensions**

Now consider the extrema of a function $f(x, y, z)$ with one constraint $C = g(x, y, z)$. Assume that (x^*, y^*, z^*) and λ^* satisfy the equations for a critical point of the Lagrangian equations

$$\nabla f_{(x^*,y^*,z^*)} = \lambda^* \nabla g_{(x^*,y^*,z^*)} \qquad \text{and} \qquad C = g(x^*, y^*, z^*). \tag{4}$$

Let

$$L(\lambda, x, y, z) = f(x, y, z) - \lambda\, [g(x, y, z) - C]$$

be the Lagrangian function for the problem. The corresponding *bordered Hessian matrix* is

$$\boxed{\mathbf{H}_4 = HL(\lambda^*, x^*, y^*, z^*) = \begin{pmatrix} 0 & -g_x & -g_y & -g_z \\ -g_x & L_{xx} & L_{xy} & L_{xz} \\ -g_y & L_{yx} & L_{yy} & L_{yz} \\ -g_z & L_{zx} & L_{zy} & L_{zz} \end{pmatrix},} \tag{5}$$

where the partial derivatives are evaluated at (x^*, y^*, z^*) and λ^*. In three dimensions, there are two directions in which we can move in the level surface, and we need two numbers to determine whether the solution of the Lagrange multiplier problem is a local maximum or local minimum. Therefore, we need to consider not only the four-by-four bordered matrix $\mathbf{H}_4$, but also a three-by-three submatrix; the submatrix is

$$\boxed{\mathbf{H}_3 = \begin{cases} \begin{pmatrix} 0 & -g_x & -g_y \\ -g_x & L_{xx} & L_{xy} \\ -g_y & L_{yx} & L_{yy} \end{pmatrix} & \text{if } g_x(x^*, y^*, z^*) \neq 0 \text{ or } g_y(x^*, y^*, z^*) \neq 0 \\[2ex] \begin{pmatrix} 0 & -g_y & -g_z \\ -g_y & L_{yy} & L_{yz} \\ -g_z & L_{zy} & L_{zz} \end{pmatrix} & \text{if } g_x(x^*, y^*, z^*) = 0 \text{ and } g_y(x^*, y^*, z^*) = 0, \end{cases}} \tag{6}$$

where the partial derivatives are evaluated at (x^*, y^*, z^*) and λ^*. Then, we have the following second derivative test.

Theorem 5. *Let f and g be real valued functions on $\mathbb{R}^3$. Let $(x^*, y^*, z^*) \in \mathbb{R}^3$ and λ^* be a solution of the Lagrange multiplier problem $\nabla f_{(x^*,y^*,z^*)} = \lambda^* \nabla g_{(x^*,y^*,z^*)}$ and $C = g(x^*, y^*, z^*)$. Assume that $\nabla g_{(x^*,y^*,z^*)} \neq \mathbf{0}$. Define the bordered Hessian matrices $\mathbf{H}_4$ and $\mathbf{H}_3$ by equations* (5) *and* (6)*.*

(a) *The point (x^*, y^*, z^*) is a local minimum of f on $c = g(x^*, y^*, z^*)$ if $-\det(\mathbf{H}_3) > 0$ and $-\det(\mathbf{H}_4) > 0$.*
(b) *The point (x^*, y^*, z^*) is a local maximum of f on $c = g(x^*, y^*, z^*)$ if $-\det(\mathbf{H}_3) < 0$ and $-\det(\mathbf{H}_4) > 0$.*
(c) *If $-\det(\mathbf{H}_4) < 0$, then the point (x^*, y^*, z^*) is a type of saddle and is neither a local minimum nor a local maximum.*

Remark 2. Again, the factor -1 in front of the determinants comes from the fact that we are considering one constraint.

Remark 3. The theorem in this case involves the determinant of a four-by-four matrix. In the cases we evaluate one of these, we expand on a row to find the answer or use the many zeroes in the matrix to get the answer in terms of the product of determinants of two submatrices. The general treatment of determinants is beyond this course and is treated in a course on linear algebra.

A way to see that the conditions on $\det(\mathbf{H}_4)$ and $\det(\mathbf{H}_3)$ are right is to take the *special case* where $g(x,y,z) = x$, $L_{xy}(\mathbf{x}^*) = 0$ $L_{xz}(\mathbf{x}^*) = 0$, and $L_{yz}(\mathbf{x}^*) = 0$. In this case,

$$\mathbf{H}_4 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & L_{xx} & 0 & 0 \\ 0 & 0 & L_{yy} & 0 \\ 0 & 0 & 0 & L_{zz} \end{pmatrix} \qquad \text{and} \qquad \mathbf{H}_3 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & L_{xx} & 0 \\ 0 & 0 & L_{yy} \end{pmatrix}.$$

Then, $-\det(\mathbf{H}_3) = L_{yy}$, and expanding $\det(\mathbf{H}_4)$ in the fourth row, $-\det(\mathbf{H}_4) = -L_{zz}\det(\mathbf{H}_3) = L_{yy}L_{zz}$. At a local minimum $L_{yy} > 0$ and $L_{zz} > 0$, so $-\det(\mathbf{H}_3) > 0$ and $-\det(\mathbf{H}_4) > 0$. Similarly, at a local maximum, $L_{yy} < 0$ and $L_{zz} < 0$, so $-\det(\mathbf{H}_3) < 0$ and $-\det(\mathbf{H}_4) > 0$. The general case takes into consideration the cross partial derivatives of L and allows the constraint function to be nonlinear. However, an argument using linear algebra reduces the result to this special case.

Example 2. Consider the problem of finding the extreme point of $f(x,y,z) = x^2 + y^2 + z^2$ on $2 = z - xy$. The method of Lagrange multipliers finds the points

$$\begin{aligned} (\lambda^*, x^*, y^*, z^*) &= (4, 0, 0, 2), \\ &= (2, 1, -1, 1), \qquad \text{and} \\ &= (2, -1, 1, 1). \end{aligned}$$

The Lagrangian function is

$$L(\lambda, x, y, z) = x^2 + y^2 + z^2 - \lambda z + \lambda xy + \lambda 2$$

with Hessian matrix

$$\mathbf{H}_4 = HL = \begin{pmatrix} 0 & y & x & -1 \\ y & 2 & \lambda & 0 \\ x & \lambda & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix}.$$

At the point $(\lambda^*, x^*, y^*, z^*, \lambda^*) = (4, 0, 0, 2)$, expanding on the first row,

$$-\det(\mathbf{H}_4) = -\det\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 2 & 4 & 0 \\ 0 & 4 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} = -\det\begin{pmatrix} 0 & 2 & 4 \\ 0 & 4 & 2 \\ -1 & 0 & 0 \end{pmatrix} = -12 < 0,$$

so the point is not a local extremum.

The calculation at the other two points is similar, so we consider the point $(\lambda^*, x^*, y^*, z^*) = (2, 1, -1, 1)$. The partial derivative $g_x(1, -1, 1) = -(-1) \neq 0$, so

$$\mathbf{H}_3 = \begin{pmatrix} 0 & y & x \\ y & 2 & \lambda \\ x & \lambda & 2 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

Expanding $\det(\mathbf{H}_3)$ on the first row,

$$\begin{aligned} -\det(\mathbf{H}_3) &= -\det\begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix} \\ &= -\det\begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} - \det\begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} \\ &= 4 + 4 = 8 > 0. \end{aligned}$$

Expanding det($\mathbf{H}_4$) on the fourth row,

$$\begin{aligned} -\det(\mathbf{H}_4) &= -\det\begin{pmatrix} 0 & -1 & 1 & -1 \\ -1 & 2 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \\ &= -\det\begin{pmatrix} -1 & 1 & -1 \\ 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix} - (2)\det\begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix} \\ &= -(0) - 2(-8) = 16 > 0. \end{aligned}$$

Thus, this point is a local minimum. A similar calculation at $(x, y, z, \lambda) = (-1, 1, 1, 2)$ shows that it is also a local minimum. When working the problem originally, we found these two points as the minima.

6. An Example with two Constraints

Example 3. Find the highest point on the set given by $x + y + z = 12$ and $z = x^2 + y^2$.

The function to be maximized is $f(x, y, z) = z$. The two constraint functions are $g_1(x, y, z) = x + y + z = 12$ and $g_2(x, y, z) = x^2 + y^2 - z = 0$.

The first order conditions are

$$\begin{aligned} f_x &= \lambda g_x + \mu h_x & \qquad 0 &= \lambda + \mu 2x \\ f_y &= \lambda g_y + \mu h_y & \qquad 0 &= \lambda + \mu 2y \\ f_z &= \lambda g_z + \mu h_z & \qquad 1 &= \lambda - \mu. \end{aligned}$$

From the third equation, we get $\lambda = 1 + \mu$. So we can eliminate this variable from the equations. They become

$$\begin{aligned} 0 &= 1 + \mu + \mu 2x \\ 0 &= 1 + \mu + \mu 2y. \end{aligned}$$

Subtracting the second from the first, we get $0 = 2\mu(x - y)$, so $\mu = 0$ or $x = y$.

Consider the first case of $\mu = 0$. This implies that $\lambda = 1$. But then, $0 = \lambda + 2x\mu = 1$, which is a contradiction. Therefore, there is no solution with $\mu = 0$.

Next, assume $y = x$. Then, from the constraints become $z = 2x^2$ and $12 = 2x + z = 2x + 2x^2$, so $0 = x^2 + x - 6 = (x + 3)(x - 2)$, and $x = 2$ or -3. If $x = 2$, then $y = 2$, $z = 2x^2 = 8$, $0 = 1 + \mu(5)$, $\mu = {}^{-1}\!/_5$, and $\lambda = 1 + \mu = {}^4\!/_5$.

If $x = y = -3$, then $z = 2x^2 = 18$, $0 = 1 + \mu(-5)$, $\mu {}^1\!/_5$, and $\lambda = 1 + \mu = {}^6\!/_5$.

We have found two critical points $(\lambda^*, \mu^*, x^*, y^*, z^*) = \left({}^4\!/_5, {}^{-1}\!/_5, 2, 2, 8, \right)$ and $\left({}^6\!/_5, {}^1\!/_5, -3, -3, 18\right)$.

For this example, $n = 3$ and $k = 2$, so $n - k = 1$. The bordered Hessian is

$$\mathbf{H}_5 = HL = \begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2x & -2y & 1 \\ -1 & -2x & -\mu 2 & 0 & 0 \\ -1 & -2y & 0 & -\mu 2 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

At $(\lambda^*, \mu^*, x^*, y^*, z^*) = \left(4/5, -1/5, 2, 2, 8, \right)$,

$$(-1)^2 \det(HL) = \det\begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -4 & -4 & 1 \\ -1 & -4 & 2/5 & 0 & 0 \\ -1 & -4 & 0 & 2/5 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix} = -\det\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ -1 & -4 & 2/5 & 0 & 0 \\ -1 & -4 & 0 & 2/5 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$$= \det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ 0 & -5 & 2/5 & 0 & 0 \\ 0 & -5 & 0 & 2/5 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} = -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 2/5 & 0 \\ 0 & -5 & 2/5 & 0 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$$= -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 2/5 & 0 \\ 0 & 0 & 2/5 & -2/5 & 0 \\ 0 & 0 & -4 & -4 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} = -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 2/5 & 0 \\ 0 & 0 & 2/5 & -2/5 & 0 \\ 0 & 0 & 0 & -8 & 1 \\ 0 & 0 & 0 & -2 & -1 \end{bmatrix}$$

$$= -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 2/5 & 0 \\ 0 & 0 & 2/5 & -2/5 & 0 \\ 0 & 0 & 0 & -8 & 1 \\ 0 & 0 & 0 & 0 & -5/4 \end{bmatrix} = 20 > 0.$$

Therefore, this point is a local minimum.

At $(\lambda^*, \mu^*, x^*, y^*, z^*) = \left(6/5, 1/5, -3, -3, 18\right)$,

$$(-1)^2 \det(HL) = \det\begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 6 & 6 & 1 \\ -1 & 6 & -2/5 & 0 & 0 \\ -1 & 6 & 0 & -2/5 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix} = -\det\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 6 & 6 & 1 \\ -1 & 6 & -2/5 & 0 & 0 \\ -1 & 6 & 0 & -2/5 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$$= \det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 6 & 6 & 1 \\ 0 & 5 & -2/5 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} = -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 5 & -2/5 & 0 & 0 \\ 0 & 0 & 6 & 6 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}$$

$$= -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 0 & -2/5 & 2/5 & 0 \\ 0 & 0 & 6 & 6 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix} = -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 0 & -2/5 & 2/5 & 0 \\ 0 & 0 & 0 & 12 & 1 \\ 0 & 0 & 0 & -2 & -1 \end{bmatrix}$$

$$= -\det\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 5 & 0 & -2/5 & 0 \\ 0 & 0 & -2/5 & 2/5 & 0 \\ 0 & 0 & 0 & 12 & 1 \\ 0 & 0 & 0 & 0 & -5/6 \end{bmatrix} = -20 < 0.$$

Therefore, this point is a local maximum.

These answers are compatible with the values of $f(x, y, z)$ at the two critical point: The first is a global minimum on the constraint set, and the second is a global maximum on the constraint set.