3

Relative Extrema for a Function

3–1. RELATIVE EXTREMA FOR A FUNCTION OF ONE VARIABLE

Let $f(x)$ be a function of x which is defined for the interval $x_1 \leqslant x \leqslant x_2$. If $f(x) - f(a) \geqslant 0$ for all values of x in the total interval $x_1 \leqslant x \leqslant x_2$, except $x = a$, we say the function has an *absolute* minimum at $x = a$. If $f(x) - f(a) > 0$ for all values of x except $x = a$ in the *subinterval*, $\alpha \leqslant x \leqslant \beta$, containing $x = a$, we say that $f(a)$ is a *relative* minimum, that is, it is a minimum with respect to all other values of $f(x)$ for the particular subinterval. Absolute and relative maxima are defined in a similar manner. The relative maximum and minimum values of a function are called relative extrema. One should note that $f(x)$ may have a number of relative extreme values in the total interval $x_1 \leqslant x \leqslant x_2$.

As an illustration, consider the function shown in Fig. 3–1. The relative extrema are $f(a), f(b), f(c), f(d)$. Using the notation introduced above, we say that $f(b)$ is a relative minimum for the interval $\alpha_b \leqslant x \leqslant \beta_b$. The absolute maximum and minimum values of f occur at $x = a$ and $x = d$, respectively.

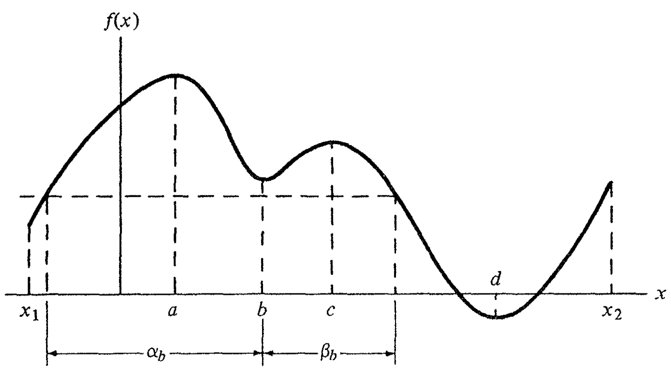


Fig. 3–1. Stationary points at points *A*, *B*, *C*, and *D*.

In general, values of x at which the slope *changes* sign correspond to relative extrema. To find the relative extrema for a continuous function, we first determine the points at which the first derivative vanishes. These points are called *stationary* points. We then test each stationary point to see if the slope *changes* sign. If the second derivative is positive (negative) the stationary point is a relative minimum (maximum). If the second derivative also vanishes, we must consider higher derivatives at the stationary point in order to determine whether the slope actually changes sign. In this case, the third derivative must also vanish for the stationary point to be a relative extremum.

Example 3–1

(1)

$$f(x) = \tfrac{1}{3}x^3 + 2x^2 + x + 5$$

Setting the first derivative equal to zero,

$$\frac{df}{dx} = x^2 + 4x + 1 = 0$$

and solving for x, we obtain

$$x_{1,2} = -2 \pm \sqrt{3}$$

The second derivative is

$$\frac{d^2f}{dx^2} = 2x + 4 = 2(x + 2)$$

Then, $x = x_1 = -2 + \sqrt{3}$ corresponds to a relative minimum and $x = x_2 = -2 - \sqrt{3}$ corresponds to a relative maximum.

(2)

$$f(x) = (x - a)^3 + c$$

The first two derivatives are

$$\frac{df}{dx} = 3(x - a)^2 \qquad \text{(a)}$$

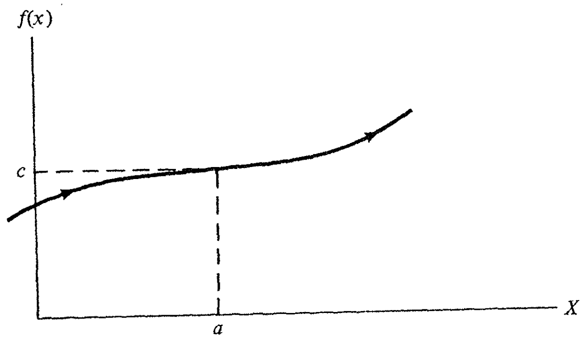
$$\frac{d^2f}{dx^2} = 6(x - a)$$

Since both derivatives vanish at $x = a$, we must consider the third derivative:

$$\frac{d^3f}{dx^3} = 6$$

The stationary point, $x = a$, is neither a relative minimum nor a relative maximum since the third derivative is finite. We could have also established this result by considering the expression for the slope. We see from (a) that the slope is positive on *both* sides of $x = a$. The general shape of this function is shown in Fig. E3–1.

Fig. E3–1



The sufficient condition for a stationary value to be a relative extremum (relative minimum (maximum) when $d^2f/dx^2 > 0$ (< 0)) follows from a consideration of the geometry of the $f(x)$ vs. x curve in the vicinity of the stationary point. We can also establish the criteria for a relative extremum from the Taylor series expansion of $f(x)$. Since this approach can be readily extended to functions of more than one independent variable we will describe it in detail.

Suppose we know the value of $f(x)$ at $x = a$ and we want $f(a + \Delta x)$ where Δx is some increment in x. If the first $n + 1$ derivatives of $f(x)$ are continuous in the interval, $a \leqslant x \leqslant a + \Delta x$, we can express $f(a + \Delta x)$ as

$$f(a + \Delta x) - f(a) = \frac{df(a)}{dx}\Delta x + \frac{1}{2}\frac{d^2f(a)}{dx^2}(\Delta x)^2 + \cdots + \frac{1}{n!}\frac{d^nf(a)}{dx^n}(\Delta x)^n + R_n \qquad (3\text{–}1)$$

where $d^jf(a)/dx^j$ denotes the jth derivative of $f(x)$ evaluated at $x = a$, and the remainder R_n is given by

$$R_n = \frac{1}{(n + 1)!}\frac{d^{n+1}f(\xi)}{dx^{n+1}}(\Delta x)^{n+1} \qquad (3\text{–}2)$$

where ξ is an unknown number between a and $a + \Delta x$. Equation (3–1) is called the Taylor series expansion* of $f(x)$ about $x = a$. If $f(x)$ is an nth-degree polynomial, the $(n + 1)$th derivative vanishes for all x and the expansion will yield the exact value of $f(a + \Delta x)$ when n terms are retained. In all other cases, there will be some error, represented by R_n, due to truncating the series at n terms. Since R_n depends on ξ, we can only establish bounds on R_n. The following example illustrates this point.

* See Ref. 1, Article 16–8.

Example 3–2

We expand $\sin x$ in a Taylor series about $x = 0$ taking $n = 2$. Using (3–1) and (3–2), and noting that $a = 0$, we obtain

$$\sin \Delta x = \Delta x + R_2 \qquad \text{(a)}$$

$$R_2 = -\frac{(\Delta x)^3}{6}\cos \xi \qquad 0 \leqslant \xi \leqslant \Delta x \qquad \text{(b)}$$

The bounds on $|R_2|$ are

$$\frac{|\Delta x|^3}{6}\cos \Delta x < |R_2| < \frac{|\Delta x|^3}{6} \qquad \text{(c)}$$

If we use (a) to find sin (0.2), the upper bound on the truncation error is $(0.2)^3/6 \approx 0.0013$.

If Δx is small with respect to unity, the first term on the right-hand side of (3–1) is the dominant term in the expansion. Also, the second term is more significant than the third, fourth, ..., nth terms. We refer to $df/dx\ \Delta x$ as the first-order increment in $f(x)$ due to the increment, Δx. Similarly, we call $\frac{1}{2}d^2f/dx^2(\Delta x)^2$ the second-order increment, and so on. Now, $f(a)$ is a relative minimum when $f(a + \Delta x) - f(a) > 0$ for all points in the neighborhood of $x = a$, that is, for all finite values of Δx in some interval, $-\eta \leqslant \Delta x \leqslant \varepsilon$, where η and ε are arbitrary small positive numbers. Considering Δx to be small, the first-order increment dominates and we can write

$$f(a + \Delta x) - f(a) = \frac{df(a)}{dx}\Delta x + (\text{second- and higher-order terms}) \qquad (3\text{–}3)$$

For $f(a + \Delta x) - f(a)$ to be positive for both positive and negative values of Δx, the first order increment must vanish, that is, $df(a)/dx$ must vanish. Note that this is a *necessary* but not sufficient condition for a relative minimum. If the first-order increment vanishes, the second-order increment will dominate:

$$f(a + \Delta x) - f(a) = \frac{1}{2}\frac{d^2f(a)}{dx^2}(\Delta x)^2 + (\text{third- and higher-order terms}) \qquad (3\text{–}4)$$

It follows from (3–4) that the second-order increment must be positive for $f(a + \Delta x) - f(a) > 0$ to be satisfied. This requires $d^2f(a)/dx^2 > 0$. Finally, the *necessary* and *sufficient* conditions for a relative minimum at $x = a$ are

$$\frac{df(a)}{dx} = 0 \qquad \frac{d^2f(a)}{dx^2} > 0 \qquad (3\text{–}5)$$

If the first two derivatives vanish at $x = a$, the third-order increment is now the dominant term in the expansion.

$$f(a + \Delta x) + f(a) = \frac{1}{6}\frac{d^3f(a)}{dx^3}(\Delta x)^3 + (\text{fourth- and higher-order terms}) \qquad (3\text{–}6)$$

Since the third-order increment depends on the sign of Δx, it must vanish for

$f(a)$ to be a relative extremum. The sufficient conditions for this case are as follows:

Relative Minimum

$$\frac{d^3f}{dx^3} = 0 \qquad \frac{d^4f}{dx^4} > 0 \tag{3-7}$$

Relative Maximum

$$\frac{d^3f}{dx^3} = 0 \qquad \frac{d^4f}{dx^4} < 0$$

The notation used in the Taylor series expansion of $f(x)$ becomes somewhat cumbersome for more than one variable. In what follows, we introduce new notation which can be readily extended to the case of n variables. First, we define Δf to be the *total* increment in $f(x)$ due to the increment, Δx.

$$\Delta f = f(x + \Delta x) - f(x) \tag{3-8}$$

This increment depends on Δx as well as x. Next, we define the differential operator, d, as

$$d(\,) = \frac{d(\,)}{dx}\Delta x \tag{3-9}$$

The result of operating on $f(x)$ with d is called the *first differential* and is denoted by df:

$$df = \frac{df(x)}{dx}\Delta x = df(x, \Delta x) \tag{3-10}$$

The first differential of $f(x)$ is a function of *two* independent variables, namely, x and Δx. If $f(x) = x$, then $df/dx = 1$ and

$$df = dx = \Delta x \tag{3-11}$$

One can use dx and Δx interchangeably; however, we will use Δx rather than dx. Higher differentials of $f(x)$ are defined by iteration. For example, the second differential is given by

$$d^2f = d(df) = \left[\frac{d}{dx}\left(\frac{df}{dx}\Delta x\right)\right]\Delta x \tag{3-12}$$

Since Δx is *independent* of x,

$$\frac{d}{dx}(\Delta x) = 0$$

and d^2f reduces to

$$d^2f = \frac{d^2f(x)}{dx^2}(\Delta x)^2 = d^2f(x, \Delta x) \tag{3-13}$$

In forming the higher differentials, we take $d(\Delta x) = 0$.

Using differential notation, the Taylor series expansion (3-1) about x can be written as

$$\Delta f = df + \frac{1}{2}d^2f + \cdots + \frac{1}{n!}d^nf + R_n \tag{3-14}$$

The first differential represents the first-order increment in $f(x)$ due to the increment, Δx. Similarly, the second differential is a measure of the second-order increment, and so on. Then, $f(x)$ is a stationary value when $df = 0$ for all permissible values of Δx. Also, the stationary point is a relative minimum (maximum) when $d^2f > 0\,(<0)$ for all permissible values of Δx. The above criteria reduce to (3-5) when the differentials are expressed in terms of the derivatives.

Rules for forming the differential of the sum or product of functions are listed below for reference. Problems 3-4 through 3-7 illustrate their application.

$f = u(x) + v(x)$

$$\begin{aligned} df &= du + dv \\ d^2f &= d(df) = d^2u + d^2v \end{aligned} \tag{3-15}$$

$f = u(x)v(x)$

$$\begin{aligned} df &= u\,dv + v\,du \\ d^2f &= u\,d^2v + 2\,du\,dv + v\,d^2u \end{aligned} \tag{3-16}$$

$f = f(y)$ where $y = y(x)$

$$\begin{aligned} df &= \frac{df}{dy}dy \\ d^2f &= \frac{d^2f}{dy^2}(dy)^2 + \frac{df}{dy}d^2y \end{aligned} \tag{3-17}$$

3-2. RELATIVE EXTREMA FOR A FUNCTION OF n INDEPENDENT VARIABLES

Let $f(x_1, x_2, \ldots, x_n)$ be a continuous function of n independent variables $(x_1, x_2, \ldots, x_n)$. We define Δf as the total increment in f due to increments in the independent variables $(\Delta x_1, \Delta x_2, \ldots, \Delta x_n)$:

$$\Delta f = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots, x_n + \Delta x_n) - f(x_1, x_2, \ldots, x_n) \tag{3-18}$$

If $\Delta f > 0\,(<0)$ for all points in the neighborhood of $(x_1, x_2, \ldots, x_n)$, we say that $f(x_1, x_2, \ldots, x_n)$ is a *relative minimum* (*maximum*). We establish criteria for a relative extremum by expanding f in an n-dimensional Taylor series. The procedure is identical to that followed in the one-dimensional case. Actually, we just have to extend the differential notation from one to n dimensions.

We define the n-dimensional differential operator as

$$d = \frac{\partial()}{\partial x_1}\Delta x_1 + \frac{\partial()}{\partial x_2}\Delta x_2 + \cdots + \frac{\partial()}{\partial x_n}\Delta x_n = \sum_{j=1}^{n}\frac{\partial()}{\partial x_j}\Delta x_i \qquad (3\text{–}19)$$

where the increments $(\Delta x_1, \Delta x_2 \cdots \Delta x_n)$ are independent of $(x_1, x_2, \ldots, x_n)$. The result obtained when d is applied to f is called the *first differential* and written as df.

$$df = \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\Delta x_j \qquad (3\text{–}20)$$

Higher differentials are defined by iteration. For example, the second differential has the form

$$d^2f = d(df) = \sum_{k=1}^{n}\frac{\partial}{\partial x_k}\left(\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\Delta x_j\right)\Delta x_k \qquad (3\text{–}21)$$

Since Δx_j are considered to be independent, (3–21) reduces to

$$d^2f = \sum_{k=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_k}\Delta x_j\,\Delta x_k \qquad (3\text{–}22)$$

Now, we let

$$\begin{aligned}\mathbf{f}^{(1)} &= \left\{\frac{\partial f}{\partial x_j}\right\}\\ \mathbf{f}^{(2)} &= \left[\frac{\partial^2 f}{\partial x_j \partial x_k}\right] \qquad j, k = 1, 2, \ldots, n\\ \Delta\mathbf{x} &= \{\Delta x_j\}\end{aligned} \qquad (3\text{–}23)$$

and the expressions for the first two differentials simplify to

$$\begin{aligned}df &= \Delta\mathbf{x}^T\mathbf{f}^{(1)}\\ d^2f &= \Delta\mathbf{x}^T\mathbf{f}^{(2)}\,\Delta\mathbf{x}\end{aligned} \qquad (3\text{–}24)$$

The Taylor series expansion for f about $(x_1, x_2, \ldots, x_n)$, when expressed in terms of differentials, has the form

$$\Delta f = df + \frac{1}{2}d^2f + \cdots + \frac{1}{n!}d^nf + R_n \qquad (3\text{–}25)$$

We say that $f(x_1, x_2, \ldots, x_n)$ is *stationary* when $df = 0$ for arbitrary $\Delta\mathbf{x}$. This requirement is satisfied only when

$$\mathbf{f}^{(1)} = \mathbf{0} \qquad (3\text{–}26)$$

Equation (3–26) represents n scalar equations, namely,

$$\frac{\partial f}{\partial x_j} = 0 \qquad j = 1, 2, \ldots, n \qquad (3\text{–}27)$$

The scalar equations corresponding to the stationary requirement are usually called the *Euler equations* for f. Note that the number of equations is equal to the number of *independent* variables.

A stationary point corresponds to a relative minimum (maximum) of f when d^2f is positive (negative) definite. It is called a neutral point when d^2f is either positive or negative semidefinite and a saddle point when d^2f is indifferent, i.e., the eigenvalues are both positive and negative. This terminology was originally introduced for the two dimensional case where it has geometrical significance.

To summarize, the solutions of the Euler equations correspond to points at which f is stationary. The classification of a stationary point is determined by the character (definite, semidefinite, indifferent) of $\mathbf{f}^{(2)}$ evaluated at the point. We are interested in the extremum problem since it is closely related to the stability problem. The extremum problem is also related to certain other problems of interest, e.g., the characteristic-value problem. In the following examples, we illustrate various special forms of f which are encountered in member system analysis.

Example 3–3

(1)

$$\begin{aligned}f &= f(y_1, y_2, \ldots, y_r)\\ y_j &= y_j(x_1, x_2, \ldots, x_n)\\ df &= \sum_{k=1}^{n}\frac{\partial f}{\partial x_k}\Delta x_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n}\frac{\partial f}{\partial y_j}\frac{\partial y_j}{\partial x_k}\right)\Delta x_k\end{aligned}$$

Now,

$$dy_j = \sum_{k=1}^{n}\frac{\partial y_j}{\partial x_k}\Delta x_k$$

It follows that

$$df = \sum_{j=1}^{r}\frac{\partial f}{\partial y_j}dy_j$$

Repeating leads to

$$d^2f = \sum_{j=1}^{r}\left[\frac{\partial f}{\partial y_j}d^2y_j + \sum_{m=1}^{r}\frac{\partial^2 f}{\partial y_m \partial y_j}dy_m\,dy_j\right]$$

(2)

Consider the double sum,

$$f = \sum_{j=1}^{s}\sum_{k=1}^{s}u_j w_{jk} v_k \qquad \text{(a)}$$

The first differential (see Prob. 3–9) has the form

$$df = \sum_{j=1}^{s}\sum_{k=1}^{s}(du_j w_{jk} v_k + u_j\,dw_{jk} v_k + u_j w_{jk}\,dv_k) \qquad \text{(b)}$$

Introducing matrix notation,

$$\mathbf{u} = \{u_j\} \qquad \mathbf{w} = [w_{jk}] \qquad \mathbf{v} = \{v_k\}$$
$$f = \mathbf{u}^T\mathbf{w}\mathbf{v} \qquad \text{(c)}$$

and letting

$$d\mathbf{u} = \{du_j\} \qquad \text{(d)}$$

and so forth, we can write df as

$$df = d(\mathbf{u}^T\mathbf{w}\mathbf{v}) = d\mathbf{u}^T\mathbf{w}\mathbf{v} + \mathbf{u}^T\, d\mathbf{w}\mathbf{v} + \mathbf{u}^T\,\mathbf{w}\,d\mathbf{v} \tag{e}$$

One operates on matrix products as if they were scalars, but the order must be preserved.

As an illustration, consider

$$f = \tfrac{1}{2}\mathbf{x}^T\mathbf{a}\mathbf{x} - \mathbf{x}^T\mathbf{c} \tag{f}$$

where $\mathbf{a}$, $\mathbf{c}$ are constant and $\mathbf{a}$ is symmetrical. Noting that $d\mathbf{a} = d\mathbf{c} = \mathbf{0}$ and $d\mathbf{x} \equiv \Delta\mathbf{x}$, the first two differentials are

$$\begin{aligned} df &= \Delta\mathbf{x}^T(\mathbf{a}\mathbf{x} - \mathbf{c}) \\ d^2f &= \Delta\mathbf{x}^T\mathbf{a}\,\Delta\mathbf{x} \end{aligned} \tag{g}$$

Comparing (g) and (3–24), we see that

$$\begin{aligned} \mathbf{f}^{(1)} &= \mathbf{a}\mathbf{x} - \mathbf{c} \\ \mathbf{f}^{(2)} &= \mathbf{a} \end{aligned} \tag{h}$$

The Euler equations are obtained by setting $\mathbf{f}^{(1)}$ equal to $\mathbf{0}$:

$$\mathbf{a}\mathbf{x} = \mathbf{c} \tag{i}$$

The solution of (i) corresponds to a stationary value of (f). If $\mathbf{a}$ is positive definite, the stationary point is a relative minimum. One can visualize the problem of solving the system $\mathbf{a}\mathbf{x} = \mathbf{c}$, where $\mathbf{a}$ is symmetrical from the point of view of finding the stationary value of a second-degree polynomial having the form $f = \frac{1}{2}\mathbf{x}^T\mathbf{a}\mathbf{x} - \mathbf{x}^T\mathbf{c}$.

(3)

Suppose $f = u/v$. Using the fact that

$$\frac{\partial}{\partial x_j}\left(\frac{u}{v}\right) = \frac{1}{v}\frac{\partial u}{\partial x_j} + u\frac{\partial}{\partial x_j}\left(\frac{1}{v}\right) = \frac{1}{v}\left(\frac{\partial u}{\partial x_j} - \frac{u}{v}\frac{\partial v}{\partial x_j}\right) \tag{a}$$

we can write

$$df = d\left(\frac{u}{v}\right) = \frac{1}{v}(du - f\,dv) \tag{b}$$

We apply (b) to

$$\lambda = \frac{\mathbf{x}^T\mathbf{a}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \tag{c}$$

where $\mathbf{a}$ is symmetrical, and obtain (see Prob. 3–5)

$$\begin{aligned} d\lambda &= \frac{2\,\Delta\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}}(\mathbf{a}\mathbf{x} - \lambda\mathbf{x}) \\ d^2\lambda &= \frac{2}{\mathbf{x}^T\mathbf{x}}(\Delta\mathbf{x}^T\mathbf{a}\,\Delta\mathbf{x} - \lambda\,\Delta\mathbf{x}^T\,\Delta\mathbf{x} - 2\,d\lambda\,\Delta\mathbf{x}^T\mathbf{x}) \end{aligned} \tag{d}$$

Setting $d\lambda = 0$ leads to the Euler equations for (c),

$$\mathbf{a}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \tag{e}$$

which we recognize as the symmetrical characteristic-value problem.

The quotient $\mathbf{x}^T\mathbf{a}\mathbf{x}/\mathbf{x}^T\mathbf{x}$, where $\mathbf{x}$ is arbitrary and $\mathbf{a}$ is symmetrical, is called *Rayleigh's quotient*. We have shown that the characteristic values of $\mathbf{a}$ are stationary values of Rayleigh's quotient. This property can be used to improve an initial estimate for a characteristic value. For a more detailed discussion, see Ref. 6 and Prob. 3–11.

3–3. LAGRANGE MULTIPLIERS

Up to this point, we have considered only the case where the function is expressed in terms of independent variables. In what follows, we discuss how one can modify the procedure to handle the case where some of the variables are not independent. This modification is conveniently effected using Lagrange multipliers.

Suppose f is expressed in terms of n variables, say $x_1, x_2, \ldots, x_n$, some of which are not independent. The general stationary requirement is

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\,dx_j = 0 \tag{3–28}$$

for all arbitrary differentials of the *independent* variables. We use dx_j instead of Δx_j to emphasize that some of the variables are dependent. In order to establish the Euler equations, we must express df in terms of the differentials of the independent variables.

Now, we suppose there are r relations between the variables, of the form

$$g_k(x_1, x_2, \ldots, x_n) = 0 \qquad k = 1, 2, \ldots, r \tag{3–29}$$

One can consider these relations as constraint conditions on the variables. Actually, there are only $n - r$ independent variables. We obtain r relations between the n differentials by operating on (3–29). Since $g_k = 0$, it follows that $dg_k = 0$. Then,

$$dg_k = \sum_{j=1}^{n} \frac{\partial g_k}{\partial x_j}\,dx_j = 0 \qquad k = 1, 2, \ldots, r \tag{3–30}$$

Using (3–30), we can express r differentials in terms of the remaining $n - r$ differentials. Finally, we reduce (3–28) to a sum involving the $n - r$ independent differentials. Equating the coefficients to zero leads to a system of $n - r$ equations which, together with the r constraint equations, are sufficient to determine the stationary points.

Example 3–4

We illustrate the procedure for $n = 2$ and $r = 1$:

$$\begin{aligned} f &= f(x_1, x_2) \\ g(x_1, x_2) &= 0 \end{aligned}$$

The first variation is

$$df = \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 \tag{a}$$

Operating on $g(x_1, x_2)$ we have

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0 \tag{b}$$

Now, we suppose $\partial g/\partial x_2 \neq 0$. Solving (b) for dx_2 (we replace dx_1 by Δx_1 to emphasize that x_1 is the independent variable.)

$$dx_2 = -\left(\frac{\partial g}{\partial x_1}\bigg/\frac{\partial g}{\partial x_2}\right)\Delta x_1 \tag{c}$$

and substituting in (a), we obtain

$$df = \left[\frac{\partial f}{\partial x_1} - \left(\frac{\partial g}{\partial x_1}\bigg/\frac{\partial g}{\partial x_2}\right)\frac{\partial f}{\partial x_2}\right]\Delta x_1 \tag{d}$$

Finally, the equations defining the stationary points are

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \left(\frac{\partial g}{\partial x_1}\bigg/\frac{\partial g}{\partial x_2}\right)\frac{\partial f}{\partial x_2} &= 0 \\ g(x_1, x_2) &= 0 \end{aligned} \tag{e}$$

To determine whether a stationary point actually corresponds to a relative extremum, we must investigate the behavior of the second differential. The general form of d^2f for a function of two variables (which are not necessarily independent) is

$$\begin{aligned} d^2f &= d\left(\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2\right) \\ &= \sum_{j=1}^{2}\sum_{k=1}^{2} \frac{\partial^2 f}{\partial x_j \partial x_k} dx_j\, dx_k + \sum_{j=1}^{2} \frac{\partial f}{\partial x_j} d^2x_j \end{aligned} \tag{f}$$

We reduce (f) to a quadratic form in the independent differential, Δx_1, using (c), and noting $d^2x_1 = 0$,

$$d^2f = (\Delta x_1)^2\left[\frac{\partial^2 f}{\partial x_1^2} + 2u\frac{\partial^2 f}{\partial x_1 \partial x_2} + u^2\frac{\partial^2 f}{\partial x_2^2} + \left(\frac{\partial u}{\partial x_1} + u\frac{\partial u}{\partial x_2}\right)\frac{\partial f}{\partial x_2}\right] \tag{g}$$

where

$$u = -\frac{\partial g}{\partial x_1}\bigg/\frac{\partial g}{\partial x_2}$$

The character of the stationary point is determined from the sign of the bracketed term.

An automatic procedure for handling constraint conditions involves the use of Lagrange multipliers. We first describe this procedure for the case of two variables and then generalize it for n variables and r restraints. The problem consists in determining the stationary values of $f(x_1, x_2)$ subject to the constraint condition, $g(x_1, x_2) = 0$. We introduce the function H, defined by

$$H(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \tag{3-31}$$

where λ is an unknown parameter, referred to as a Lagrange multiplier. We

consider x_1, x_2 and λ to be *independent* variables, and require H to be stationary. The Euler equations for H are

$$\begin{aligned} \frac{\partial H}{\partial x_1} &= \frac{\partial f}{\partial x_1} + \lambda\frac{\partial g}{\partial x_1} = 0 \\ \frac{\partial H}{\partial x_2} &= \frac{\partial f}{\partial x_2} + \lambda\frac{\partial g}{\partial x_2} = 0 \\ \frac{\partial H}{\partial \lambda} &= g(x_1, x_2) = 0 \end{aligned} \tag{3-32}$$

We suppose $\partial g/\partial x_2 \neq 0$. Then, solving the second equation in (3-32) for λ, and substituting in the first equation, we obtain

$$\lambda = -\frac{\partial f}{\partial x_2}\bigg/\frac{\partial g}{\partial x_2} \tag{3-33}$$

and

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \left(\frac{\partial g}{\partial x_1}\bigg/\frac{\partial g}{\partial x_2}\right)\frac{\partial f}{\partial x_2} &= 0 \\ g(x_1, x_2) &= 0 \end{aligned} \tag{3-34}$$

Equations (3-34) and (e) of the previous example are identical. We see that the Euler equations for H are the stationary conditions for f *including* the effect of constraints.

Example 3-5

$$\begin{aligned} f &= 3x_1^2 + 2x_2^2 + 2x_1 + 7x_2 \\ g &= x_1 - x_2 = 0 \end{aligned}$$

We form $H = f + \lambda g$,

$$H = 3x_1^2 + 2x_2^2 + 2x_1 + 7x_2 + \lambda(x_1 - x_2)$$

The stationary requirement for H treating x_1, x_2, and λ as independent variables is

$$\begin{aligned} 6x_1 + 2 + \lambda &= 0 \\ 4x_2 + 7 - \lambda &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$$

Solving this system for x_1, x_2 and λ we obtain

$$\begin{aligned} \lambda &= 4x_2 + 7 \\ x_1 &= x_2 = -9/10 \end{aligned}$$

This procedure can be readily generalized to the case of n variables and r constraints. The problem consists of determining the stationary values of $f(x_1, x_2, \ldots, x_n)$, subject to the constraints $g_k(x_1, x_2, \ldots, x_n) = 0$, where $k = 1, 2, \ldots, r$. There will be r Lagrange multipliers for this case, and H has the form

$$H = f + \sum_{k=1}^{r} \lambda_k g_k = H(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_r) \tag{3-35}$$

The Euler equations for H are

$$\frac{\partial f}{\partial x_i} + \sum_{k=1}^{r} \lambda_k \frac{\partial g_k}{\partial x_i} = 0 \qquad i = 1, 2, \ldots, n \tag{3-36}$$

$$g_k = 0 \qquad k = 1, 2, \ldots, r \tag{3-37}$$

We first solve r equations in (3-36) for the r Lagrange multipliers, and then determine the n coordinates of the stationary points from the remaining $n - r$ equations in (3-36) and the r constraint equations (3-37). The use of Lagrange multipliers to introduce constraint conditions usually reduces the amount of algebra.

REFERENCES

1. THOMAS, G. B., JR., *Calculus and Analytical Geometry*, Addison-Wesley Publishing Co., Reading, Mass., 1953.
2. COURANT; R., Differential and Integral Calculus, Vol. 1, Blackie, London, 1937.
3. COURANT, R., Differential and Integral Calculus, Vol. 2, Interscience Publishers, New York, 1936.
4. HANCOCK, H., *Theory of Maxima and Minima*, Dover Publications, New York, 1960.
5. APOSTOL, T. M., *Mathematical Analysis*, Addison-Wesley Publishing Co., Reading, Mass., 1957.
6. CRANDALL, S. H., *Engineering Analysis*, McGraw-Hill, New York, 1956.
7. HILDEBRAND, F. B., *Methods of Applied Mathematics*, Prentice-Hall, New York, 1952.

PROBLEMS

3-1. Determine the relative extrema for
(a) $f(x) = 2x^2 + 4x + 5$
(b) $f(x) = -2x^2 + 8x + 10$
(c) $f(x) = ax^2 + 2bx + c$
(d) $f(x) = x^3 + 2x^2 + x + 10$
(e) $f(x) = \frac{1}{3}x^3 + 2x^2 + 4x + 15$
(f) $f(x) = (x - a)^4 + (x - a)^2$
(g) $f(x) = \frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx + d$

3-2. Expand $\cos x$ in a Taylor series about $x = 0$, taking $n = 3$. Determine the upper and lower bounds on R_3.

3-3. Expand $(1 + x)^{1/2}$ in a Taylor series about $x = 0$ taking $n = 2$. Determine upper and lower bounds on R_2.

3-4. Find df and d^2f for
(a) $f = x^2 + 2x + 5$
(b) $f = 3x^3 + 2x^2 + 5x + 6$
(c) $f = x^2 \sin x$
(d) $f = \cos y$ where $y = x^3$

3-5. Let $f = u(x)/v(x)$. Show that

$$df = \frac{1}{v}(du - f\,dv)$$

$$d^2f = \frac{1}{v}(d^2u - fd^2v) - 2\frac{dv\,df}{v}$$

3-6. Let u_1, u_2, u_3 be functions of x and $f = f(u_1, u_2, u_3)$. Determine df.

3-7. Suppose $f = u(x)w(y)$ where $y = y(x)$. Determine expressions for df and d^2f. Apply to
(a) $u = x^3 - x$
(b) $w = \cos y$
(c) $y = x^2$

3-8. Find the first two differentials for the following functions:
(a) $f = x_1^3 + 3x_1^2x_2^2 + x_1x_2^3$
(b) $f = 3x_1^2 + 6x_1x_2 + 9x_2^2 + 5x_1 - 4x_2$

3-9. Consider $f = uv$, where

$$u = u(y_1, y_2) \qquad v = v(y_1, y_2)$$

and

$$y_1 = y_1(x_1, x_2) \qquad y_2 = y_2(x_1, x_2)$$

Show that

$$df = d(uv) = u\,dv + v\,du \qquad d^2f = ud^2v + 2\,du\,dv + vd^2u$$

Note that the rule for forming the differential of a product is independent of whether the terms are functions of the independent variables (x_1, x_2) or of dependent variables.

3-10. Classify the stationary points for the following functions:
(a) $f = 3x_1^3 + 3x_2^2 - 9x_1 + 12x_2 - 10$
(b) $f = 3x_1^2 + 6x_1x_2 + 2x_2^2 + 2x_1 + 7x_2$
(c) $f = 3x_1^2 + 6x_1x_2 + 3x_2^2 + 2x_1 + 2x_2$
(d) $f = 3x_1^2 + 6x_1x_2 + 4x_2^2 + 2x_1 + 7x_2$
(e) $f = 3x_1^3 + 6x_1x_2 + 3x_2^2 - 3x_1$

3-11. Consider Rayleigh's quotient,

$$\lambda = \frac{\mathbf{x}^T\mathbf{a}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{x}^T\mathbf{a}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

where $\mathbf{x}$ is arbitrary. Since $\mathbf{a}$ is symmetrical, its characteristic vectors are linearly independent and we can express $\mathbf{x}$ as

$$\mathbf{x} = \sum_{j=1}^{n} c_j\mathbf{Q}_j$$

where $\mathbf{Q}_j$ $(j = 1, 2, \ldots, n)$ are the normalized characteristic vectors for $\mathbf{a}$.
(a) Show that

$$\lambda = \frac{\sum_{j=1}^{n} \lambda_j c_j^2}{\sum_{j=1}^{n} c_j^2}$$

(b) Suppose $\mathbf{x}$ differs only slightly from $\mathbf{Q}_k$. Then, $|c_j| \ll |c_k|$ for $j \neq k$. Specialize (a) for this case. *Hint*: Factor out λ_k and c_k^2.

(c) Use (b) to obtain an improved estimate for λ.

$$\mathbf{a} = \begin{bmatrix} 3 & 1 \\ 1 & \frac{3}{2} \end{bmatrix}$$

$$\mathbf{x} \approx \{1, -3\}$$

The exact result is

$$\lambda = 1 \qquad \mathbf{x} = \{1, -2\}$$

3–12. Using Lagrange multipliers, determine the stationary values for the following constrained functions:

(a) $f = x_1^2 - x_2^2$
$g = x_1^2 + x_2 = 0$

(b) $f = x_1^2 + x_2^2 + x_3^2$
$g_1 = x_1 + x_2 + x_3 - 1 = 0$
$g_2 = x_1 - x_2 + 2x_3 + 2 = 0$

3–13. Consider the problem of finding the stationary values of $f = \mathbf{x}^T\mathbf{a}\mathbf{x} = \mathbf{x}^T\mathbf{a}^T\mathbf{x}$ subject to the constraint condition, $\mathbf{x}^T\mathbf{x} = 1$. Using (3–36) we write

$$H = f + \lambda g = \mathbf{x}^T\mathbf{a}\mathbf{x} - \lambda(\mathbf{x}^T\mathbf{x} - 1)$$

(a) Show that the equations defining the stationary points of f are

$$\mathbf{a}\mathbf{x} = \lambda\mathbf{x} \qquad \mathbf{x}^T\mathbf{x} = 1$$

(b) Relate this problem to the characteristic value problem for a symmetrical matrix.

3–14. Suppose $f = \mathbf{x}^T\mathbf{x}$ and $g = 1 - \mathbf{x}^T\mathbf{a}\mathbf{x} = 0$ where $\mathbf{a}^T = \mathbf{a}$. Show that the Euler equations for H have the form

$$\mathbf{a}\mathbf{x} = \frac{1}{\lambda}\mathbf{x} \qquad \mathbf{x}^T\mathbf{a}\mathbf{x} = 1$$

We see that the Lagrange multipliers are the reciprocals of the characteristic values of $\mathbf{a}$. How are the multipliers related to the stationary values of f?

4

Differential Geometry of a Member Element

The geometry of a member element is defined once the curve corresponding to the reference axis and the properties of the normal cross section (such as area, moments of inertia, etc.) are specified. In this chapter, we first discuss the differential geometry of a space curve in considerable detail and then extend the results to a member element. Our primary objective is to introduce the concept of a local reference frame for a member.

4–1. PARAMETRIC REPRESENTATION OF A SPACE CURVE

A curve is defined as the locus of points whose position vector* is a function of a single parameter. We take an orthogonal cartesian reference frame having directions X_1, X_2, and X_3 (see Fig. 4–1). Let $\bar{r}$ be the position vector to a point

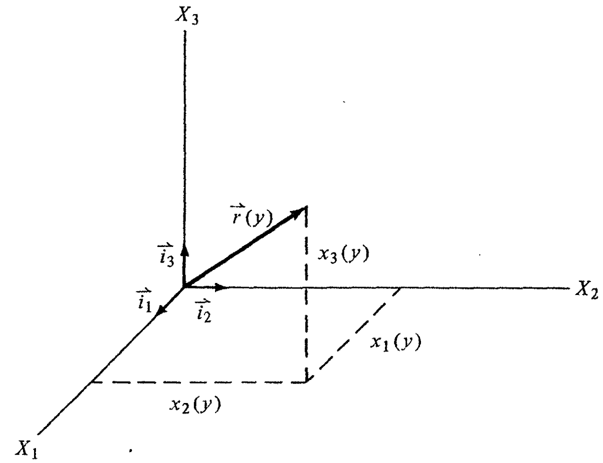


Fig. 4–1. Cartesian reference frame with position vector $\bar{r}(y)$.

* The vector directed from the origin of a fixed reference frame to a point is called the *position vector*. A knowledge of vectors is assumed. For a review, see Ref. 1.