REVIEW OF CALCULUS

Herman J. Bierens

Pennsylvania State University

(January 28, 2004)

1. Summation

Let $x_1, x_2, ..., x_n$ be a sequence of numbers. The sum of these numbers is usually denoted by

$$x_1 + x_2 + ... + x_n = \sum_{j=1}^{n} x_j, \quad \text{or} \quad x_1 + x_2 + ... + x_n = \sum_{j=1}^{n} x_j .$$

The index "*j*" may be replaced by any other variable name. Thus, $x_1 + x_2 + ... + x_n = \sum_{i=1}^{n} x_i$ as well.

Next, let $y_1, y_2, ..., y_m$ be another sequence of numbers. Then we can write

$$(x_1 + x_2 + ... + x_n)(y_1 + y_2 + ... + y_n) = (x_1 + x_2 + ... + x_n)\left(\sum_{j=1}^{m} y_j\right)$$

$$= x_1\left(\sum_{j=1}^{m} y_j\right) + x_2\left(\sum_{j=1}^{m} y_j\right) + ... + x_n\left(\sum_{j=1}^{m} y_j\right) \tag{1}$$

$$= \sum_{j=1}^{m} x_1 y_j + \sum_{j=1}^{m} x_2 y_j + ... + \sum_{j=1}^{m} x_n y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j .$$

Of course, the order of the summation does not matter:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j = \sum_{j=1}^{m}\sum_{i=1}^{n} x_i y_j .$$

Moreover, replacing $y_1, y_2, ..., y_m$ in (1) by $x_1, x_2, ..., x_n$ it follows that

$$\left(\sum_{j=1}^{n} x_j\right)^2 = (x_1 + x_2 + ... + x_n)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j . \tag{2}$$

Note that the reason for using different indices *i* and *j* is that the summation in (2) is done in two steps. First, for each index *i* we sum up $x_i x_j$ for *j* = 1,...,*n*, and then we sum up $x_i \sum_{j=1}^{n} x_j$ for *i* =

1,....,n. The same applies to (1).

The average of the numbers $x_1, x_2, \ldots, x_n$ is usually denoted by

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{j=1}^{n} x_i.$$

In particular, we have

$$\begin{aligned}\sum_{j=1}^{n}(x_j - \bar{x}) &= \sum_{j=1}^{n} x_j - \sum_{j=1}^{n}\bar{x} = \sum_{j=1}^{n} x_j - n.\bar{x} \\ &= \sum_{j=1}^{n} x_j - n.\frac{1}{n}\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} x_j = 0.\end{aligned} \qquad (3)$$

This easy result will prove useful in regression analysis.

2. Special sums

Consider the sum $1 + 2 + 3 + \ldots + n = \sum_{j=1}^{n} j$. There is an easy formula for this sum:

$$\sum_{j=1}^{n} j = n(n+1)/2. \qquad (4)$$

Rather than memorizing this formula, it is better to memorize how it is derived. Since the order of the summation does not matter, we can write $\sum_{j=1}^{n} j$ in two ways:

$$\begin{aligned}\sum_{j=1}^{n} j &= 1 + 2 + 3 + \ldots + n, \quad \textit{and} \\ \sum_{j=1}^{n} j &= n + n-1 + n-2 + \ldots + 1\end{aligned} \qquad (5)$$

Adding up the left and right-hand sides of the two equations in (5) yields

$$2\sum_{j=1}^{n} j = (n+1) + (n+1) + (n+1) + \ldots + (n+1) = n(n+1). \qquad (6)$$

Dividing (6) by 2, the result (4) follows.

Next, consider the sum

$$1 + x + x^2 + x^3 + \ldots + x^n = x^0 + x^1 + x^2 + x^3 + \ldots + x^n = \sum_{j=0}^{n} x^j,$$

where x is any number. The formula for this sum is:

$$\sum_{j=0}^{n} x^j = \frac{1 - x^{n+1}}{1 - x}. \tag{7}$$

Again, it is easier to memorize how this formula is derived than to memorize the formula itself. First, observe that $\sum_{j=0}^{n} x^j = 1 + \sum_{j=1}^{n} x^j$. Next, replace the index j by i-1. Then we can write:

$$\begin{aligned}\sum_{j=0}^{n} x^j &= 1 + \sum_{j=1}^{n} x^j = 1 + \sum_{i=0}^{n-1} x^{i+1} = 1 + x.\sum_{i=0}^{n-1} x^i = 1 + x.\left(\sum_{j=0}^{n-1} x^j\right) \\ &= 1 + x.\left(\sum_{j=0}^{n} x^j - x^n\right) = 1 - x^{n+1} + x.\sum_{j=0}^{n} x^j.\end{aligned} \tag{8}$$

Solving equation (8) for $\sum_{j=0}^{n} x^j$ yields the formula (7).

3. Limits

The limit of a sequence y_n, $n = 1,2,3,....$, of numbers is a number y such that y_n approaches y if n increases to infinity. The formal definition of a limit is:

> *A sequence y_n, $n = 1,2,3,....$, of numbers has a limit y, say, denoted by $y = \lim_{n\to\infty} y_n$, if and only if for every number $\varepsilon > 0$ there exists an index n_0 (which may depend on ε) such that $|y_n - y| < \varepsilon$ for all $n \geq n_0$.*

For example, if $y_n = 1/n$ then $\lim_{n\to\infty} y_n = 0$. To see this, pick an arbitrary $\varepsilon > 0$. Then $|y_n| = 1/n < \varepsilon$ if $n > 1/\varepsilon$. Thus in this case the index n_0 is the smallest natural number $\geq 1/\varepsilon$. Another example is the case $y_n = x^n$, where $|x| < 1$. Given such a number x and an arbitrary number $\varepsilon > 0$, it is possible to find a positive integer n_0 such that $|x|^{n_0} < \varepsilon$, and then $|y_n| = |x^n| = |x|^n < \varepsilon$ for all $n \geq n_0$. Hence

$$\lim_{n\to\infty} x^n = 0 \ \ if \ |x| < 1. \tag{9}$$

Consequently, it follows from (7) and (9) that:

$$\lim_{n\to\infty}\sum_{j=0}^{n} x^j = \frac{1 - \lim_{n\to\infty}x^{n+1}}{1 - x} = \frac{1 - x.\lim_{n\to\infty}x^n}{1 - x} = \frac{1}{1 - x} \;\; if \;\; |x| < 1. \tag{10}$$

The left-hand side of this equation is usually denoted by $\sum_{j=0}^{\infty}x^j$. Thus,

$$\sum_{j=0}^{\infty} x^j = \frac{1}{1 - x} \;\; if \;\; |x| < 1. \tag{11}$$

More generally, if for a sequence $x_0, x_1, x_2, ..., x_n, ...$ of numbers, $\lim_{n\to\infty}\sum_{j=0}^{n}x_j$ exists, then this limit is denoted by $\sum_{j=0}^{\infty}x_j$. Moreover, it is left as an exercise to verify that

$$If \;\; \lim_{n\to\infty}\sum_{j=0}^{n} x_j \;\; exists \;\; then \;\; \lim_{m\to\infty}\sum_{j=m}^{\infty} x_j = 0. \tag{12}$$

Not every sequence has a limit, though. For example, if $y_n = (-1)^n$ then the limit does not exist, because then $|y_n| = |-1|^n = 1^n = 1$ for n = 1,2,3,...., so that for 0 < ε < 1 and all positive natural numbers n, $|y_n| > \varepsilon$.

4. Functions

4.1 Linear and quadratic functions and their roots

A real function $f(x)$ assigns a real number $y = f(x)$ to x. Important classes of functions are the linear functions:

$$f(x) = \alpha + \beta.x \tag{13}$$

and the quadratic functions:

$$f(x) = \alpha + \beta.x + \gamma.x^2, \tag{14}$$

where α, β and γ are given constants.

The roots of a function $f(x)$ are the values of x for which $f(x) = 0$. In the linear case (13) there is only one root, namely $x = -\alpha/\beta$, provided that $\beta \neq 0$. In the quadratic case (14) the number of roots is either 0, 1 or 2, depending on what α, β and γ are. In order to derive these roots, observe first that $(x + a)^2 = x^2 + 2.a.x + a^2$. Next, assume that $\gamma \neq 0$. Then

$$\alpha + \beta.x + \gamma.x^2 = 0 \Rightarrow x^2 + \frac{\beta}{\gamma}.x = -\frac{\alpha}{\gamma} \Rightarrow x^2 + 2.\frac{\beta}{2\gamma}.x = -\frac{\alpha}{\gamma}$$

$$\Rightarrow x^2 + 2.\frac{\beta}{2\gamma}.x + \left(\frac{\beta}{2\gamma}\right)^2 = -\frac{\alpha}{\gamma} + \left(\frac{\beta}{2\gamma}\right)^2 \tag{15}$$

$$\Rightarrow \left(x + \frac{\beta}{2\gamma}\right)^2 = \frac{\beta^2 - 4\alpha\gamma}{4\gamma^2}.$$

If $\beta^2 - 4\alpha\gamma > 0$ then the last equality in (15) implies that the quadratic function (14) has two roots:

$$x_1 = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma}, \quad x_2 = \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma}. \tag{16}$$

If $\beta^2 - 4\alpha = 0$ then the quadratic function (14) has only one root:

$$x = \frac{-\beta}{2\gamma}, \tag{17}$$

and if $\beta^2 - 4\alpha < 0$ the quadratic function (14) has no real roots.[1]

4.2 The exp(.) *and* ln(.) *functions*

The exponential function exp(x) is defined as

$$\exp(x) = e^x, \ where \ e \approx 2.7182818285 \tag{18}$$

It can be shown[2] that

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \tag{19}$$

---

1 However, the roots are then complex-valued:

$$x_1 = \frac{-\beta - \boldsymbol{i}.\sqrt{4\alpha\gamma - \beta^2}}{2\gamma}, \quad x_2 = \frac{-\beta + \boldsymbol{i}.\sqrt{4\alpha\gamma - \beta^2}}{2\gamma}, \text{ where } \boldsymbol{i} = \sqrt{-1}.$$

2 But that requires advanced calculus!

where $k!$ (read: k factorial) is the product of the natural numbers 1 to k: $k! = 1 \times 2 \times 3 \times ... \times k$, and 0! is defined as 1: $0! = 1$.

The numerical value of the number e in (18) is only an approximation, though. The true number e has an infinite number of decimal digits without a repeating pattern, and is therefore irrational. We can only express e exactly as a limit:

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n \to \infty} \sum_{k=0}^{n} \frac{1}{k!}.$$

The natural logarithm is the inverse of the exponential function:

The natural logarithm, $\ln(x)$, *is for each* $x > 0$ *a number* y *such that* $\exp(y) = x$.

Since $\exp(y) > 0$, the function $\ln(x)$ is only defined for $x > 0$. Important properties of the function $\ln(x)$ are:

$$\begin{aligned}
&\ln(x) < 0 \;\; \textit{for} \;\; 0 < x < 1, \\
&\ln(x) = 0 \;\; \textit{for} \;\; x = 1, \\
&\ln(x) > 0 \;\; \textit{for} \;\; x > 1, \\
&\ln(x.y) = \ln(x) + \ln(y) \;\; \textit{for} \;\; x > 0 \;\; \textit{and} \;\; y > 0, \\
&\ln(x/y) = \ln(x) - \ln(y) \;\; \textit{for} \;\; x > 0 \;\; \textit{and} \;\; y > 0, \\
&\ln(x^{\alpha}) = \alpha.\ln(x) \;\; \textit{for} \;\; x > 0 \;\; \textit{and any} \;\; \alpha.
\end{aligned} \tag{20}$$

It is left as an exercise to verify these properties from the definition of $\ln(x)$.

4.3 Continuity

Continuity of a function $f(x)$ in a point x_0 can be defined in two (equivalent) ways. The first way is via limits:

A function $f(x)$ *is continuous in* x_0 *if and only if* $f(x_0) = \lim_{n \to \infty} f(x_0 + 1/n)$ *and* $f(x_0) = \lim_{n \to \infty} f(x_0 - 1/n)$.

This definition is equivalent to the following official definition:

A function f(x) is continuous in x_0 *if and only if for each* $\varepsilon > 0$ *there exists a* $\delta > 0$ (*possibly depending on* x_0) *such that* $|f(x) - f(x_0)| < \varepsilon$ *if* $|x - x_0| < \delta$.

It is not hard to show that the second definition implies the first one. The proof that the other way around is also true is harder and therefore omitted.

The second definition gives rise to another definition of a limit:

$\lim_{x \to x_0} f(x) = y$ *if and only if for each* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that* $|f(x) - y| < \varepsilon$ *if* $|x - x_0| < \delta$.

Thus, *f*(*x*) is continuous in x_0 if $\lim_{x \to x_0} f(x) = f(x_0)$.

For example, the linear function (13) is continuous in all *x*, and so is the quadratic function (14). The latter can be show as follows. Let x_0 be arbitrary and fixed, and let $|x - x_0| < 1$. Then

$$
\begin{aligned}
|f(x) - f(x_0)| &= |\beta(x - x_0) + \gamma(x^2 - x_0^2)| = |\beta(x - x_0) + \gamma(x - x_0)(x + x_0)| \\
&= |\beta(x - x_0) + \gamma(x - x_0)^2 + 2x_0\gamma(x - x_0)| \qquad (21) \\
&\leq |\beta||x - x_0| + |\gamma||x - x_0|^2 + 2|x_0\gamma||x - x_0| < (|\beta| + |\gamma| + 2|x_0\gamma|)|x - x_0|,
\end{aligned}
$$

where the last inequality in (21) follows from the fact that $|x - x_0|^2 < |x - x_0|$ if $|x - x_0| < 1$. Next, choose an arbitrary $\varepsilon > 0$, and let $\delta = \min[1, \varepsilon/(|\beta| + |\gamma| + 2|x_0\gamma|)]$. Then it follows from (21) that $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$.

Also the exponential function (18) is continuous in all *x*, and the ln(*x*) function is continuous in all $x > 0$.

5. Derivatives

5.1 What is a derivative?

The derivative of a function $f(x)$ is denoted by $f'(x)$, and is defined by

$$f'(x) = \lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{\delta}. \tag{22}$$

An alternative notation for a derivative is $df(x)/dx$. In order for a derivative to exist, the limit in (22) must exist and be the same regardless whether $\delta > 0$ (so that $\delta \downarrow 0$) or $\delta < 0$ (so that $\delta \uparrow 0$). If so, the function involved is said to be *differentiable* in *x*.

The derivation of $f'(x)$ is illustrated in the following figure.

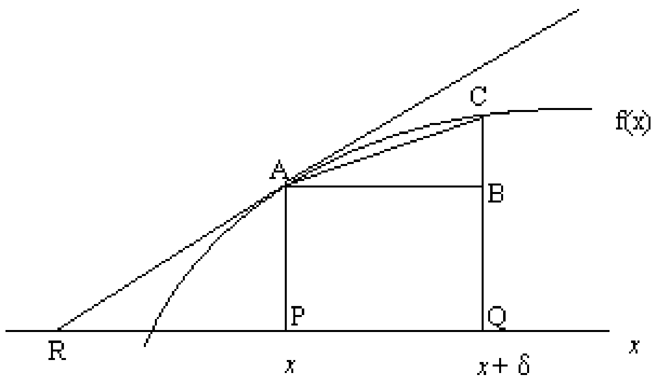

Figure 1: Derivative

The curved line in Figure 1 represents the function $f(x)$. The length of the line piece between the points P and A, $\overline{P \to A}$ say, is the value of the function $f(x)$ in x at point P, and $\overline{Q \to C}$ is equal to $f(x+\delta)$. Moreover, $\overline{P \to Q} = \overline{A \to B} = \delta$. Thus,

$$\frac{f(x+\delta) - f(x)}{\delta} = \frac{\overline{C \to B}}{\overline{A \to B}},$$

which is the tangents[3] of the angle between the legs $A \to B$ and $A \to C$ of the triangle ABC. Now if we let $\delta \to 0$ then this angle approaches the angle of the line through the points R and A with the

[3] The tangents of an angle φ is defined as: $\tan(\varphi) = \sin(\varphi)/\cos(\varphi)$.

horizontal axis. This line is called the *tangent line* of $f(x)$ in point A. Thus, the derivative of $f(x)$ in x at point P is:

$$f'(x) = \frac{\overline{A{\to}P}}{\overline{R{\to}P}}.$$

This ratio is called the *slope* of the function $f(x)$ in x at point P.

The quadratic function (14) is differentiable in every x:

$$f(x) = \alpha + \beta.x + \gamma.x^2 \Rightarrow f'(x) = \beta + 2\gamma.x. \tag{23}$$

To see this, observe that in the case (14),

$$\begin{aligned} &\lim_{\delta\to 0} \frac{f(x+\delta) - f(x)}{\delta} = \lim_{\delta\to 0} \frac{\beta.\delta + \gamma.(x+\delta)^2 - \gamma.x^2}{\delta} \\ &= \lim_{\delta\to 0} \frac{\beta.\delta + 2\gamma.\delta x + \gamma\delta^2}{\delta} = \lim_{\delta\to 0} (\beta + 2\gamma x + \gamma\delta) = \beta + 2\gamma x. \end{aligned} \tag{24}$$

Of course, the derivative of the linear function (13) follows from (23) by setting $\gamma = 0$.

The exponential function has itself as derivative:

$$f(x) = e^x \Rightarrow f'(x) = e^x, \tag{25}$$

because it follows from (19) that

$$\begin{aligned} &\lim_{\delta\to 0} \frac{f(x+\delta) - f(x)}{\delta} = \lim_{\delta\to 0} \frac{e^{x+\delta} - e^x}{\delta} = e^x.\lim_{\delta\to 0} \frac{e^\delta - 1}{\delta} = e^x.\lim_{\delta\to 0} \frac{\sum_{k=0}^{\infty}\delta^k/k! - 1}{\delta} \\ &= e^x.\lim_{\delta\to 0} \frac{1 + \sum_{k=1}^{\infty}\delta^k/k! - 1}{\delta} = e^x.\lim_{\delta\to 0} \sum_{k=1}^{\infty}\delta^{k-1}/k! = e^x.\lim_{\delta\to 0} \sum_{k=0}^{\infty}\delta^k/(k+1)! \\ &= e^x.\left(1 + \lim_{\delta\to 0} \sum_{k=1}^{\infty}\delta^k/(k+1)!\right) = e^x. \end{aligned} \tag{26}$$

Moreover, the derivative of $\ln(x)$ is $1/x$:

$$f(x) = \ln(x) \Rightarrow f'(x) = 1/x, \ \textit{for all } x > 0, \tag{27}$$

To prove (27), let $y = \ln(x)$. Then $x = \exp(y)$, hence it follows from (25) that

$$dx/dy = d\exp(y)/dy = \exp(y) = x. \tag{28}$$

Taking the reciprocal of (28) it follows that

$$dy/dx = d\ln(x)/dx = 1/x. \tag{29}$$

It should be noted that there are many function that are not differentiable in some points, even functions that are continuous in every point. For example, the absolute value function $f(x) = |x|$ is continuous in every x, but is not differentiable in $x = 0$.

5.2 The chain rule

Next, consider two differentiable functions $f(x)$ and $g(x)$, and let $h(x) = f(g(x))$. The question is: Given the derivatives $f'(x)$ and $g'(x)$, what is the derivative $h'(x) = df(g(x))/dx$? The answer is the *chain rule*:

$$h'(x) = \frac{df(g(x))}{dx} = \frac{df(g(x))}{dg(x)} \times \frac{dg(x)}{dx} = f'(g(x)).g'(x). \tag{30}$$

For example, let $f(x) = \ln(x)$ and $g(x) = x^2+1$. Then $h(x) = f(g(x)) = \ln(x^2+1)$, $f'(x) = 1/x$, and $g'(x) = 2x$, hence it follows from (30) that

$$h'(x) = \frac{2x}{x^2+1}.$$

The proof of (30) is left as an exercise.

6. Integrals

6.1 What is an integral?

The concept of an integral is illustrated in Figure 2:

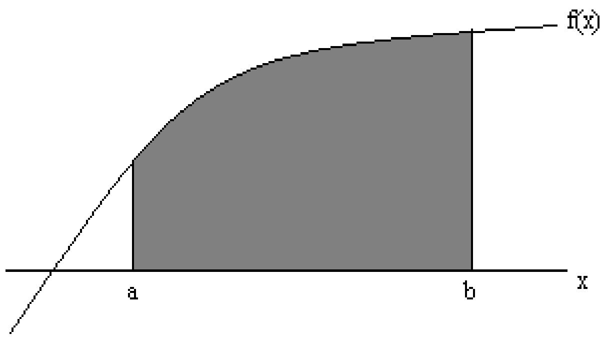

Figure 2: The integral $\int_a^b f(x)dx$ = grey area.

The integral of a function $f(x)$ over an interval $[a,b]$, denoted by $\int_a^b f(x)dx$, is the grey area in Figure 2. This area can be (roughly) approximated by a sum[4] of rectangle areas with height $f(x)$ and width dx, as illustrated in Figure 3:

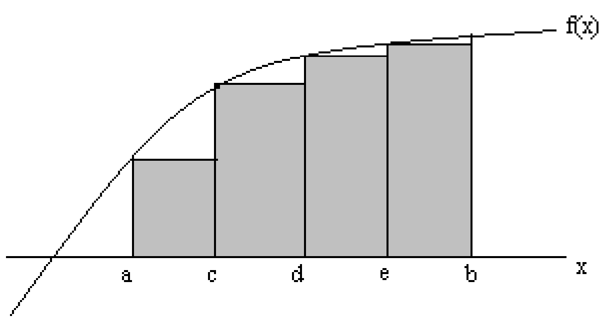

Figure 3: Approximation of $\int_a^b f(x)dx$

The first rectangle has area $f(a)\times(c-a)$, the second has area $f(c)\times(d-c)$, the third has area

4 Therefore, the integral symbol $\int$ is actually a stylized version of the letter *S* in "*S*um".

$f(d)\times(e-d)$, and the last one has area $f(e)\times(d-b)$. Assuming that the intervals $[a,c]$, $[c,d]$, $[d,e]$ and $[e,b]$ have equal length $(b - a)/4 = dx$, say, so that

$$c = a + (b-a)/4,\ d = a + 2(b-a)/4,\ e + 3(b-a)/4,$$

the total grey area in Figure 3 is:

$$\sum_{k=0}^{n-1} f(a + k(b-a)/n)\times(b-a)/n = \sum_{\substack{x=a+k(b-a)/n \\ dx=(b-a)/n \\ k=0,1,\ldots,n-1}} f(x)dx,\ for\ n = 4.$$

Letting $n \to \infty$, the limit of this sum is the grey area in Figure 2: $\int_a^b f(x)dx$.

Note that if $f(x) < 0$ on $[a,b]$ then $\int_a^b f(x)dx < 0$, as follows by flipping the pictures in Figures 2 and 3 vertically 180 degrees. Therefore, the integral of $f(x)$ on $[a,b]$ in Figure 4 below is the *difference* of the darker grey area above the horizontal axis, and the lighter grey area below the horizontal axis.

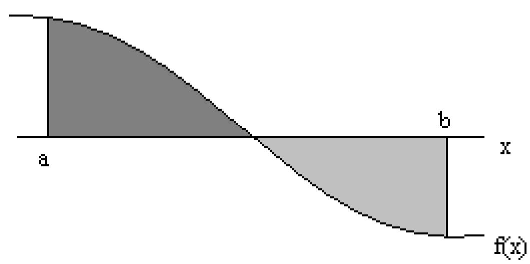

Figure 4: Integral $\int_a^b f(x)dx$ if $f(x)$ flips sign

Moreover,

$$\int_a^b f(x)dx = -\int_b^a f(x)dx. \tag{31}$$

because the latter integral should be interpreted as the grey area in Figure 1 looking from the *back* (the negative side) of the picture: If we flip Figure 1 horizontally 180 degrees, then point b will be at the left of point a, and the new viewpoint is now *behind* Figure 1:

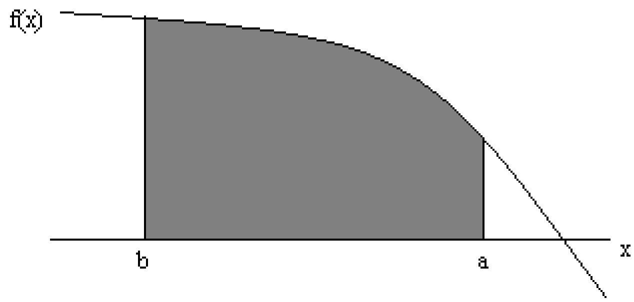

Figure 5: Back side of Figure 1

6.2 Derivative of an integral

Let point *e* in Figure 3 be $b - \delta$, where $\delta > 0$ is very small. Then the approximation

$$\int_{b-\delta}^{b} f(x)dx \approx f(b-\delta)\delta \approx f(b)\delta$$

will be close, and so will be

$$\frac{\int_{b-\delta}^{b} f(x)dx}{\delta} \approx f(b).$$

Therefore,

$$\lim_{\delta \downarrow 0} \frac{\int_{b-\delta}^{b} f(x)dx}{\delta} = f(b).$$

Similarly it follows that

$$\lim_{\delta \downarrow 0} \frac{\int_{b}^{b+\delta} f(x)dx}{\delta} = f(b). \tag{32}$$

More generally, we have:

$$\textit{Let } F(x) = \int_{a}^{x} f(u)du, \textit{ where } x > a. \textit{ Then } F'(x) = f(x). \tag{33}$$

In order to verify this, observe that

$$F^{\prime}(x) = \lim_{\delta \to 0} \frac{F(x+\delta) - F(x)}{\delta} = \lim_{\delta \to 0} \frac{\int_a^{x+\delta} f(u)du - \int_a^x f(u)du}{\delta} = \lim_{\delta \to 0} \frac{\int_x^{x+\delta} f(u)du}{\delta} = f(x), \tag{34}$$

where the last equality follows from (32).

Moreover, it follows from (31):

$$\textit{Let } F(x) = \int_x^b f(u)du, \textit{ where } x < b. \textit{ Then } F^{\prime}(x) = -f(x). \tag{35}$$

The proof of (35) is left as an exercise.

7. *Functions of two variables, and their partial derivatives*

A real function $f(x,y)$ with two arguments, x and y, assigns a real number $z = f(x,y)$ to a pair (x,y). These functions are called *bivariate* functions. For example, consider the bivariate quadratic function

$$f(x,y) = x^2 + y^2. \tag{36}$$

The shape of this function is a hyperbola:

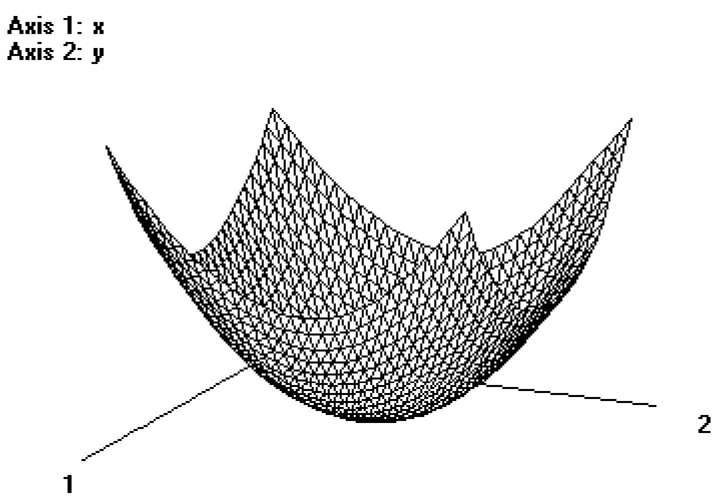

Figure 6: The function $f(x,y) = x^2 + y^2$ on the square $-1 < x < 1,\ -1 < y < 1$.

If the function $f(x,y)$ is differentiable in both arguments, then we can take the derivative of $f(x,y)$ to x, *treating y as a constant*. This derivative is called the *partial derivative of* $f(x,y)$ *to x*, and is denoted by $\partial f(x,y)/\partial x$ or $\frac{\partial f(x,y)}{\partial x}$:

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\delta \to 0} \frac{f(x+\delta,y) - f(x,y)}{\delta}. \tag{37}$$

Similarly, we can also take the derivative of $f(x,y)$ to y, *treating x as a constant*. This derivative is called the *partial derivative of* $f(x,y)$ *to y*, and is denoted by $\partial f(x,y)/\partial y$ or $\frac{\partial f(x,y)}{\partial y}$:

$$\frac{\partial f(x,y)}{\partial y} = \lim_{\delta \to 0} \frac{f(x,y+\delta) - f(x,y)}{\delta}. \tag{38}$$

For example, in the case (36) we have $\partial f(x,y)/\partial x = 2x$, $\partial f(x,y)/\partial y = 2y$.

The general bivariate quadratic function takes the form

$$f(x,y) = \gamma_1(x - \beta_1 y - \alpha_1)^2 + \gamma_2(y - \beta_2 x - \alpha_2)^2, \tag{39}$$

where $\alpha_1,\beta_1,\gamma_1,\alpha_2,\beta_2,\gamma_2$ are constants. Then

$$\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= 2\gamma_1(x - \beta_1 y - \alpha_1) - 2\gamma_2\beta_2(y - \beta_2 x - \alpha_2) \\ \frac{\partial f(x,y)}{\partial y} &= 2\gamma_2(y - \beta_2 x - \alpha_2) - 2\gamma_1\beta_1(x - \beta_1 y - \alpha_1) \end{aligned} \tag{40}$$

8. The minimum or maximum of a bivariate quadratic function

Clearly, the function (36) is minimal zero for $x = 0$ and $y = 0$. In this point the partial derivatives involved are zero. This can easily be verified directly, but also from Figure 5: In the point (0,0) the hyperbola touches the horizontal plane at zero level, and is above zero level for any other point (x,y). Thus, the point (x,y) for which the function $f(x,y) = x^2 + y^2$ is minimal can be found by solving the so-called *first-order conditions*:

$$\partial f(x,y)/\partial x = 2x = 0,$$
$$\partial f(x,y)/\partial y = 2y = 0.$$

Next, let us have a look at the general bivariate quadratic function (39). If $\gamma_1 > 0$ and $\gamma_2 > 0$ then (39) can be written as

$$f(x,y) = \left(\sqrt{\gamma_1}x - \beta_1\sqrt{\gamma_1}y - \alpha_1\sqrt{\gamma_1}\right)^2 + \left(\sqrt{\gamma_2}y - \beta_2\sqrt{\gamma_2}x - \alpha_2\sqrt{\gamma_2}\right)^2. \tag{41}$$

The shape of this function is similar to Figure 5, except that the hyperbola involved will be shifted, squeezed and/or turned horizontally. Also, it is clear that this function is minimal if x and y are such that

$$\begin{aligned} &\sqrt{\gamma_1}x - \beta_1\sqrt{\gamma_1}y - \alpha_1\sqrt{\gamma_1} = 0 \Rightarrow x - \beta_1 y - \alpha_1 = 0 \\ &\sqrt{\gamma_2}y - \beta_2\sqrt{\gamma_2}x - \alpha_2\sqrt{\gamma_2} = 0 \Rightarrow y - \beta_2 x - \alpha_2 = 0 \end{aligned} \tag{42}$$

However, the same conditions can be derived by setting the partial derivatives (40) equal to zero:

$$\left.\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= 2\gamma_1(x - \beta_1 y - \alpha_1) - 2\gamma_2\beta_2(y - \beta_2 x - \alpha_2) = 0 \\ \frac{\partial f(x,y)}{\partial y} &= 2\gamma_2(y - \beta_2 x - \alpha_2) - 2\gamma_1\beta_1(x - \beta_1 y - \alpha_1) = 0 \end{aligned}\right\}$$
$$\Rightarrow \left\{\begin{aligned} x - \beta_1 y - \alpha_1 = 0 \\ y - \beta_2 x - \alpha_2 = 0 \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} x = (\alpha_1 + \alpha_2\beta_1)/(1-\beta_1\beta_2) \\ y = (\alpha_2 + \alpha_1\beta_2)/(1-\beta_1\beta_2) \end{aligned}\right. \tag{43}$$

provided that $\beta_1\beta_2 \neq 1$.

If $\gamma_1 < 0$ and $\gamma_2 < 0$ then (39) can be written as

$$f(x,y) = -\left(\sqrt{-\gamma_1}x - \beta_1\sqrt{-\gamma_1}y - \alpha_1\sqrt{-\gamma_1}\right)^2 - \left(\sqrt{-\gamma_2}y - \beta_2\sqrt{-\gamma_2}x - \alpha_2\sqrt{-\gamma_2}\right)^2. \tag{44}$$

It is easy to verify that now the bivariate quadratic function (39) takes a maximum, and that the point (x,y) where it is maximal can be obtained by solving the first-order conditions (43), provided that $\beta_1\beta_2 \neq 1$.

9. Exercises

1. Prove (12).
2. Prove (20).
3. Prove that $\exp(x)$ is continuous in all x.
4. Prove that $\ln(x)$ is continuous in all $x > 0$, using one or more of the properties (20).
5. Prove the chain rule (30). Hint: Write $g(x+\delta) = g(x) + (g(x+\delta) - g(x))$, and use the fact that by the continuity of $g(x)$ in x, $\lim_{\delta \to 0}(g(x+\delta) - g(x)) = 0$.
6. Prove (35) by modifying (34) to this case.
7. Determine the set of points (x,y) where the bivariate quadratic function (39) is minimal or maximal, for the case $\beta_1\beta_2 = 1$ and $\gamma_1 > \gamma_2/\beta_1$.