MA 105 : Calculus
Division 3, Lecture 26

Prof. Sudhir R. Ghorpade
IIT Bombay

Recap of the previous lecture

- Discriminant Test for functions of two variables
- Examples
- Closed sets and boundary points
- Finding global extrema: Extreme Value Theorem
- Critical points and global extrema
- Recipe for finding global extrema. Example
- Orthogonal gradient theorem. Example

Recall the last result that we have discussed.

Proposition (Orthogonal Gradient Theorem)

Let $E \subset \mathbb{R}^2$ and (x_0, y_0) be an interior point of E. Let C be a parametrized curve in E passing through (x_0, y_0) given by $(x(t), y(t))$, where $x, y : [\alpha, \beta] \to E$, and let $t_0 \in (\alpha, \beta)$ with $x(t_0) = x_0$, $y(t_0) = y_0$. Suppose a function $f : E \to \mathbb{R}$, when restricted to C, has a local extremum at (x_0, y_0). If x and y are differentiable at t_0, and f is differentiable at (x_0, y_0), then

$$(\nabla f)(x_0, y_0) \cdot (x'(t_0), y'(t_0)) = 0,$$

that is, the gradient of f at (x_0, y_0) is perpendicular to the tangent vector of C at (x_0, y_0).

Proof. Define $\phi : [\alpha, \beta] \to \mathbb{R}$ by $\phi(t) := f(x(t), y(t))$. Then ϕ has a local extremum at t_0, and so $\phi'(t_0) = 0$. But by the chain rule (ii), $\phi'(t_0) = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0)$. □

Constrained Extrema

Let $D \subset \mathbb{R}^2$, and let $f, g : D \to \mathbb{R}$. Let us address the question of finding the maximum and the minimum of f on D subject to the constraint $g = 0$. The equation $g(x, y) = 0$, $(x, y) \in D$ often defines a curve C in D implicitly.

For example, let $D := \{(x, y) \in \mathbb{R}^2 : 1/4 \leq x^2 + y^2 \leq 4\}$, and let $g(x, y) := x^2 + y^2 - 1$ for $(x, y) \in D$. Then the curve C defined by $g(x, y) = 0$ is the unit circle in $\mathbb{R}^2$, and it lies in D.

Let (x_0, y_0) be an interior point of D such that $g(x_0, y_0) = 0$. Suppose f has a local extremum at (x_0, y_0) subject to the constraint $g(x, y) = 0$. We would like to show that the gradient vectors $(\nabla f)(x_0, y_0)$ and $(\nabla g)(x_0, y_0)$ are parallel.

First we give a general idea, and then we shall quote a precise result. Suppose f and g are differentiable at (x_0, y_0).

Assume that we are able to solve the equation $g(x, y) = 0$ for y in terms of x near x_0, that is, there a function η defined near x_0 such that $g(x, \eta(x)) = 0$ and $\eta(x_0) = y_0$. If $\eta'(x_0)$ exists, then $g_x(x_0, y_0) + g_y(x_0, y_0)\eta'(x_0) = 0$ by the chain rule (ii).

Consider the function $\phi(x) := f(x, \eta(x))$ for x near x_0. Now ϕ has a local extremum at x_0, and so $\phi'(x_0) = 0$, that is, $f_x(x_0, y_0) + f_y(x_0, y_0)\,\eta'(x_0) = 0$ again by the chain rule (ii).

It follows that $f_x(x_0, y_0)g_y(x_0, y_0) = f_y(x_0, y_0)g_x(x_0, y_0)$, that is, the gradient vectors $(\nabla f)(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0))$ and $(\nabla g)(x_0, y_0) = (g_x(x_0, y_0), g_y(x_0, y_0))$ are parallel.

In fact, if $g_y(x_0, y_0) \neq 0$, then $(\nabla f)(x_0, y_0) = \lambda_0(\nabla g)(x_0, y_0)$, where $\lambda_0 := f_y(x_0, y_0)/g_y(x_0, y_0)$.

Similarly, if $g(x, y) = 0$ can be solved for x in terms of y near y_0 and if $g_x(x_0, y_0) \neq 0$, then $(\nabla f)(x_0, y_0) = \lambda_0(\nabla g)(x_0, y_0)$, where $\lambda_0 := f_x(x_0, y_0)/g_x(x_0, y_0)$.

Proposition (Lagrange Multiplier Theorem)

Let $D \subset \mathbb{R}^2$, and let (x_0, y_0) be an interior point of D. Suppose f, $g : D \to \mathbb{R}$ have continuous partial derivatives in a neighbourhood of (x_0, y_0). Let $C := \{(x, y) \in D : g(x, y) = 0\}$. Suppose (i) $g(x_0, y_0) = 0$, (ii) $(\nabla g)(x_0, y_0) \neq (0, 0)$, and (iii) the function f, when restricted to C, has a local extremum at (x_0, y_0). Then there is $\lambda_0 \in \mathbb{R}$ such that

$$(\nabla f)(x_0, y_0) = \lambda_0 (\nabla g)(x_0, y_0).$$

The real number λ_0 is called a **Lagrange multiplier**.

We now have the following procedure to determine the global extremum of a function $f : D \to \mathbb{R}$, subject to the constraint $g = 0$. Consider a new variable λ, called an **undetermined multiplier**, and seek simultaneous solutions of the following 3 equations in the 3 unknowns x, y, λ:
$(\nabla f)(x, y) = \lambda (\nabla g)(x, y)$ and $g(x, y) = 0$, $(x, y) \in D$.

Suppose f has a global extremum on the set $\{(x, y) \in D : g(x, y) = 0\}$. (For example, when this set is nonempty, closed and bounded, and f is continuous on it). Then it is also a local extremum of f, and so it is attained either at a simultaneous solution (x_0, y_0) of the above two equations where $\nabla g(x_0, y_0) \neq (0, 0)$, or at a point (x_1, y_1) where $g(x_1, y_1) = 0$ and $(\nabla g)(x_1, y_1) = (0, 0)$.

Examples: (i) Consider the problem of maximizing/minimizing $f(x, y) := x\,y$ subject to $x + y = 2$. For this purpose, we find the global extrema of $\phi(x) := f(x, 2 - x) = x(2 - x)$, $x \in \mathbb{R}$. Since $\phi'(x) = 2(1 - x)$, $x \in \mathbb{R}$, we see that $\phi'(x) > 0$ if $x < 1$ and $\phi'(x) < 0$ if $x > 1$. Also, ϕ is continuous at 1. As in the first derivative test, $\phi(1) = 1$ is the global maximum of ϕ, that is, 1 is the constrained maximum of f. Also, $\phi(x) \to -\infty$ as $x \to \pm\infty$, and so f has no constrained minimum.

(ii) Consider the problem of maximizing/minimizing $f(x, y) := x\,y$, subject to $x^2 + y^2 - 1 = 0$.

Let $g(x, y) := x^2 + y^2 - 1$ for $(x, y) \in \mathbb{R}^2$. Note that the set $\{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$, that is, the unit circle, is nonempty, closed and bounded, and f is continuous on it. Now $(\nabla f)(x, y) = \lambda\,(\nabla g)(x, y)$ and $g(x, y) = 0$ means

$$y = 2\lambda x, \quad x = 2\lambda y, \quad \text{and} \quad x^2 + y^2 - 1 = 0.$$

Then $y\,x = 4\lambda^2 xy$, that is, $4\lambda^2 = 1$, since $x\,y \neq 0$. Thus $\lambda = \pm 1/2$, and the simultaneous solutions of the above equations are given by $(x, y) = \left(\pm 1/\sqrt{2},\, \pm 1/\sqrt{2}\right)$. Also, $(\nabla g)(x, y) \neq (0, 0)$ whenever $g(x, y) = 0$.

Thus the hypotheses of the Lagrange Multiplier Theorem are satisfied. Hence the maximum of f on the unit circle is $f\left(1/\sqrt{2},\, 1/\sqrt{2}\right) = f(-1/\sqrt{2},\, -1/\sqrt{2}) = 1/2$, while the minimum of f on the unit circle is $f\left(1/\sqrt{2},\, -1/\sqrt{2}\right) = f\left(-1/\sqrt{2},\, 1/\sqrt{2}\right) = -1/2$.

Lagrange Multipliers for Functions of 3 Variables

The Lagrange Multiplier Theorem can be adapted to a subset D of $\mathbb{R}^3$ and functions $f, g : D \to \mathbb{R}$. If f and g have continuous partial derivatives, then a global extremum of f on the set $\{(x, y, z) \in D : g(x, y, z) = 0\}$ can be found by comparing the values of f at the simultaneous solutions of the following 4 equations in the 4 unknowns x, y, z, λ:
$\nabla f = \lambda \nabla g$ and $g = 0$ at which $\nabla g \neq \mathbf{0}$. We must also consider the values of f at points where $g = 0$ and $\nabla g = \mathbf{0}$.

Further, the result can be adapted to a situation where there are two constraints, namely $g = 0$ and $h = 0$. In this case, we compare the values of f at the simultaneous solutions of the following 5 equations in the 5 unknowns x, y, z, λ, μ:
$\nabla f = \lambda \nabla g + \mu \nabla h$ and $g = 0 = h$ at which $\nabla g \neq \mathbf{0}, \nabla h \neq \mathbf{0}$ and $\nabla g \nparallel \nabla h$. We must also consider the values of f at points where $g = 0 = h$, and $\nabla g = \mathbf{0}$ or $\nabla h = \mathbf{0}$ or $\nabla g \parallel \nabla h$.

Multiple constraints

Explanation: Let $D \subset \mathbb{R}^3$ and $f : D \to \mathbb{R}$. Let C denote the curve in D obtained by intersecting the surfaces S_g and S_h defined by $g = 0$ and $h = 0$ respectively. Suppose f, when restricted to C, has a local extremum at a point P_0 of C, and the tangent vector to the curve C at P_0 is nonzero. By the three-dimensional analogue of the orthogonal gradient theorem, this tangent vector is perpendicular to $(\nabla f)(P_0)$

Suppose $(\nabla g)(P_0) \neq \mathbf{0}$, $(\nabla h)(P_0) \neq \mathbf{0}$ and $\nabla g \not\parallel \nabla h$. Now the normal lines to the surfaces S_g and S_h at P_0 are perpendicular to the tangent vector to the curve C at P_0, since C lies in S_g as well as in S_h. But the plane in $\mathbb{R}^3$ perpendicular to the tangent vector to the curve C at P_0 is two-dimensional, and so there are $\lambda_0, \mu_0 \in \mathbb{R}$ such that

$$(\nabla f)(P_0) = \lambda_0(\nabla g)(P_0) + \mu_0(\nabla h)(P_0).$$

Examples:

(i) Let us find the maximum and the minimum values of the function $f(x, y, z) = x + 2y + 3z$ on the curve of intersection of the plane $x - y + z = 1$ and the cylinder $x^2 + y^2 = 1$.

The continuous function f attains its bounds on the nonempty, closed and bounded subset $\{(x, y, z) : x - y + z = 1 = x^2 + y^2\}$.

The undetermined multipliers $\lambda, \mu \in \mathbb{R}$ satisfy

$$(1, 2, 3) = \lambda\,(1, -1, 1) + \mu\,(2x, 2y, 0),$$

that is, $\lambda + 2\mu x = 1, \quad -\lambda + 2\mu y = 2$ and $\lambda = 3$. This gives

$$\mu x = -1, \quad \mu y = 5/2, \quad \text{and so} \quad \mu^2 = (-1)^2 + (5/2)^2.$$

Hence $\mu = \pm\sqrt{29}/2$, $x = \mp 2/\sqrt{29}$ and $y = \pm 5/\sqrt{29}$, and so $z = 1 - x + y = 1 \pm 7/\sqrt{29}$. Also, $(1, -1, 1) \not\parallel (2x, 2y, 0)$.

It follows that the constrained maximum of f is $3 + \sqrt{29}$ and the constrained minimum of f is $3 - \sqrt{29}$.

(ii) Let us determine a point on the intersection of the planes $x+y+z=1$ and $3x+2y+z=6$ that is closest to the origin.

Define $f(x,y,z) := x^2+y^2+z^2$ for $(x,y,z) \in \mathbb{R}^3$. We wish to find a point $(x,y,z) \in \mathbb{R}^3$ at which f attains its minimum subject to constraints $g(x,y,z) = x+y+z-1 = 0$ and $h(x,y,z) = 3x+2y+z-6 = 0$.

The set $E := \{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0 = h(x,y,z)\}$ is nonempty since $(3,-1,-1) \in E$. It is closed since g and h are continuous, but it is not bounded. Let

$$E_1 := \{(x,y,z) \in E : \|(x,y,z)\| \leq \|(3,-1,-1)\|\}.$$

Then the set E_1 is nonempty, closed and bounded, and so the continuous function attains its minimum on E_1. Further, the minimum of f on E_1 is equal to the minimum of f on E. Thus f attains its minimum on E.

The equations $\nabla f = \lambda \nabla g + \mu \nabla h$ yield

$$x = \frac{\lambda + 3\mu}{2}, \quad y = \frac{\lambda + 2\mu}{2}, \quad z = \frac{\lambda + \mu}{2}.$$

Since $g(x, y, z) = 0 = h(x, y, z)$, we obtain $3\lambda + 6\mu = 2$ and $3\lambda + 7\mu = 6$. This gives $\lambda = -22/3$ and $\mu = 4$. Hence $P_0 := (x_0, y_0, z_0) = (7/3, 1/3, -5/3)$.

We note that $(\nabla g)(P_0) = (1, 1, 1)$ and $(\nabla h)(P_0) = (3, 2, 1)$ are nonzero and nonparallel. Hence f attains its constrained minimum on E at P_0.

We remark that f is not bounded above on the straight line E, and so it cannot have a constrained maximum on E.

ALITER:

$x + y + z = 1$ and $3x + 2y + z = 6 \implies 1 - x - y = z = 6 - 3x - 2y \implies y = 5 - 2x$ and $z = -4 + x$. Hence $x^2 + y^2 + z^2 = x^2 + (5 - 2x)^2 + (-4 + x)^2 = 6x^2 - 28x + 41$.

Therefore our problem reduces to minimizing the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) := 6x^2 - 28x + 41$, $x \in \mathbb{R}$.
Now $f'(x) = 4(3x - 7)$ for $x \in \mathbb{R}$. Clearly,
$x < 7/3 \implies f'(x) < 0 \implies f$ is decreasing on $(-\infty, 7/3)$,
$x > 7/3 \implies f'(x) > 0 \implies f$ is increasing on $(7/3, \infty)$.
Also, f is continuous at $7/3$. As in the First Derivative Test, f has a global minimum at $x = 7/3$.
Further, $x = 7/3 \implies y = 1/3$, $z = -5/3$.
Thus the point on the intersection of the planes that is closest to the origin is $(7/3, 1/3, -5/3)$.

‘Lagrange’ on the Eiffel Tower

‘Lagrange’ is among the 72 names inscribed on Eiffel Tower.





Double Integral on a Rectangle

The concept of the Riemann integral of a function was motivated by our attempt to find 'the area under a curve'. We now look for a concept which will let us find 'the volume under a surface'. We shall assume that the **volume** of a **cuboid** $[a,b]\times[c,d]\times[p,q]$ is equal to $(b-a)(d-c)(q-p)$.

Let $R := [a,b]\times[c,d]$ be a rectangle in $\mathbb{R}^2$ with $a<b$, $c<d$. Let $f : R \to \mathbb{R}$ be a bounded function. Define

$$m(f) := \inf\{f(x,y) : (x,y) \in R\},\ M(f) := \sup\{f(x,y) : (x,y) \in R\}.$$

Let $n, k \in \mathbb{N}$, and consider a **partition** P of R given by $P := \{(x_i, y_j) : i = 0,1,\ldots,n \text{ and } j = 0,1,\ldots,k\}$, where $a := x_0 < x_1 < \cdots < x_n := b,\ \ c = y_0 < y_1 < \cdots < y_k = d$.

The points in P divide the rectangle R into nk nonoverlapping subrectangles $[x_{i-1}, x_i]\times[y_{j-1}, y_j]$, $i = 1,\ldots,n$; $j = 1,\ldots,k$.

Define

$$\begin{aligned} m_{i,j}(f) &:= \inf\{f(x,y) : (x,y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\}, \\ M_{i,j}(f) &:= \sup\{f(x,y) : (x,y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\}. \end{aligned}$$

Clearly, $m(f) \leq m_{i,j}(f) \leq M_{i,j}(f) \leq M(f)$ for $i = 1, \ldots, n$; $j = 1, \ldots, k$.

Define the **lower double sum** and the **upper double sum** of f with respect to P by

$$\begin{aligned} L(P,f) &:= \sum_{i=1}^{n} \sum_{j=1}^{k} m_{i,j}(f)(x_i - x_{i-1})(y_j - y_{j-1}), \\ U(P,f) &:= \sum_{i=1}^{n} \sum_{j=1}^{k} M_{i,j}(f)(x_i - x_{i-1})(y_j - y_{j-1}). \end{aligned}$$

Since $\sum_{i=1}^{n}(x_i - x_{i-1}) = b - a$ and $\sum_{j=1}^{k}(y_j - y_{j-1}) = d - c$, we obtain

$$m(f)(b-a)(d-c) \leq L(P,f) \leq U(P,f) \leq M(f)(b-a)(d-c).$$

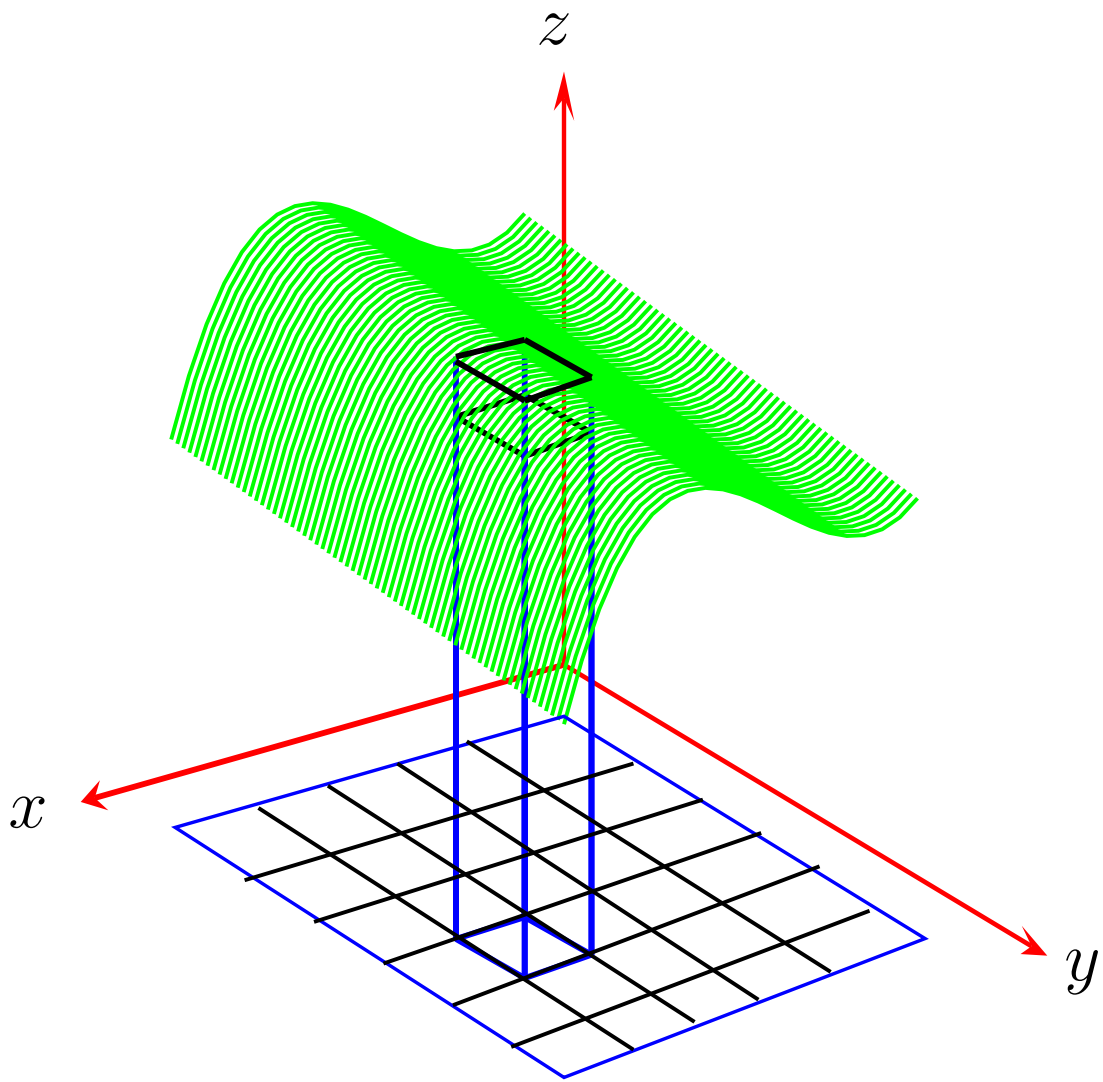

Figure: Summands of lower and upper double sums.

Let P_1 and P_2 be partitions of the rectangle R. A lower double sum increases and a upper double sum decreases when a partition is refined. Let P^* be a common refinement of P_1 and P_2. Then $L(P_1, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_2, f)$. Define

$$\begin{aligned} L(f) &:= \sup\{L(P, f) : P \text{ is a partition of } R\}, \\ U(f) &:= \inf\{U(P, f) : P \text{ is a partition of } R\}. \end{aligned}$$

$L(f)$ is called the **lower double integral** of f and $U(f)$ is called the **upper double integral** of f.

Then $L(f) \leq U(f)$ by using the definitions of sup and inf.

Definition

A bounded function $f : R \to \mathbb{R}$ is said to be **(double) integrable** on R if $L(f) = U(f)$.

In this case, the **double integral** of f on R is the common value $U(f) = L(f)$, and it is denoted by

$$\iint_R f \quad \text{or} \quad \iint_R f(x,y)d(x,y).$$

Let $f : R \to \mathbb{R}$ be integrable and nonnegative. The double integral of f on R gives the volume of the solid $E_f := \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in R \text{ and } 0 \le z \le f(x,y)\}$ under the surface $z = f(x,y)$ and above the rectangle R.

Examples:

(i) Let $f(x,y) := 1$ for all $(x,y) \in R$. Then $L(P,f) = (b-a)(d-c) = U(P,f)$ for every partition P of R. Hence f is integrable on R, and its double integral on R is equal to $(b-a)(d-c)$.

(ii) Define the **bivariate Dirichlet function** $f : R \to \mathbb{R}$ by

$$f(x,y) := \begin{cases} 1 & \text{if } x \text{ and } y \text{ are rational numbers,} \\ 0 & \text{if } x \text{ or } y \text{ is an irrational number.} \end{cases}$$

Then f is a bounded function on R. For each partition P of R,

$$m_{i,j}(f) = 0 \text{ and } M_{i,j}(f) = 1, \quad i = 1, \ldots, n;\ j = 1, \ldots, k,$$

and so $L(P, f) = 0$ and $U(P, f) = (b - a)(d - c)$. Thus $L(f) = 0$ and $U(f) = (b - a)(d - c)$. Since $L(f) \neq U(f)$, f is not integrable.

(iii) Let $\phi : [a, b] \to \mathbb{R}$ be bounded, and define $f : R \to \mathbb{R}$ by $f(x, y) := \phi(x)$ for $(x, y) \in R$. Then f is a bounded function. Let $P := \{(x_i, y_j) : i = 0, 1, \ldots, n \text{ and } j = 0, 1, \ldots, k\}$ be any partition of R. Then $P_1 := \{x_0, x_1, \ldots, x_n\}$ is a partition of $[a, b]$, and $m_{i,j}(f) = m_i(\phi)$, $i = 1, \ldots, n$; $j = 1, \ldots, k$, and so

$$\sum_{i=1}^{n}\sum_{j=1}^{n} m_{i,j}(f)(x_i - x_{i-1})(y_j - y_{j-1}) = (d-c)\sum_{i=1}^{n} m_i(\phi)(x_i - x_{i-1}).$$

Thus $L(P, f) = (d-c)L(P_1, \phi)$. Also, $U(P, f) = (d-c)U(P_1, \phi)$.

Further, if $Q := \{x_0, x_1, \ldots, x_n\}$ is a partition of $[a, b]$, then $Q = P_1$, where $P := \{(x_i, y_j) : i = 0, 1, \ldots, n,\ j = 0, 1\}$. So

$L(f) = (d - c)L(\phi)$ and $U(f) = (d - c)U(\phi)$. Hence

f is integrable on $R \iff \phi$ is integrable on $[a, b]$, and in this case, the double integral of f on R is equal to $(d - c)$ times the Riemann integral of ϕ on $[a, b]$.

The following result gives a necessary and sufficient condition for the integrability of a bounded function on R.

Theorem (Riemann condition)

Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be a bounded function. Then f is integrable if and only if for every $\epsilon > 0$, there is a partition P_ϵ of $[a, b] \times [c, d]$ such that $U(P_\epsilon, f) - L(P_\epsilon, f) < \epsilon$.

The proof is very similar to the proof of the result about the Riemann condition in the one variable case.