# Vv255 Lecture 14

Dr Jing Liu

UM-SJTU Joint Institute

June 22, 2017

- Problems like the post office problem is known as constrained optimization

$$\begin{aligned} \text{max} \qquad & V = xyz \\ \text{subject to} \qquad & x + 2y + 2z = 108 \end{aligned}$$

- The way we used is to solve the constraint equation for one of the variables in terms of the others and make the substitution.

$$\begin{aligned} x + 2y + 2z = 108 \implies \qquad & x = 108 - 2y - 2z \\ \implies \qquad & V = (108 - 2y - 2z)yz = 108yz - 2y^2z - 2yz^2 \end{aligned}$$

- This converts it into an unconstrained optimization of maximizing

$$\text{max} \qquad V = 108yz - 2y^2z - 2yz^2$$

Q: Why is this approach inadequate for some constrained optimizations?

- Even if we can isolate one variable in the constrained equation, the method of solving constrained optimization by substitution do not always work.

- Consider how to find the points in $\mathbb{R}^3$ on the surface

$$x^2 - y^2 - 1 = 0$$

that are closest to the origin.

- This is a constrained optimization problem

$$\begin{array}{rl} \text{min} & f(x,y,z) = x^2 + y^2 + z^2 \\ \text{subject to} & g(x,y,z) = x^2 - y^2 - 1 = 0 \end{array}$$

- If make $y^2$ the subject, we have

$$y^2 = x^2 - 1$$

- To find the points we need to minimize the following

$$h(x,z) = x^2 + (x^2 - 1) + z^2 = 2x^2 + z^2 - 1$$

- Using basic geometry, we expect the local minimum is the global minimum,

$$h(x, z) = 2x^2 + z^2 - 1$$

$$\implies \nabla h = \begin{bmatrix} 4x \\ 2z \end{bmatrix} = \mathbf{0}$$

$$\implies x = 0 \quad \text{and} \quad z = 0$$

Q: Is this the local minimum, and therefore the global minimum?

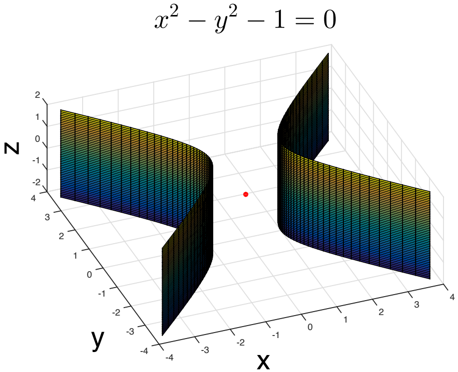
- But there is no point on the cylinder $x^2 - y^2 - 1 = 0$ where $x = 0$,

$$x^2 = 1 + y^2$$

$$\implies x^2 \geq 1$$

$$\implies x \neq 0$$

Q: What went wrong here?




- What happened was that we have found a minimum in the domain of

$$h(x, z) = 2x^2 + z^2 - 1 \qquad \text{where} \quad (x, z) \in \mathbb{R}^2$$

which is actually NOT the same as the set of points on the cylinder.

- A better way to handle constrains is known as the Lagrange multiplier.

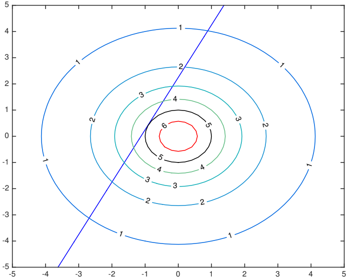
- Suppose that we are trying to

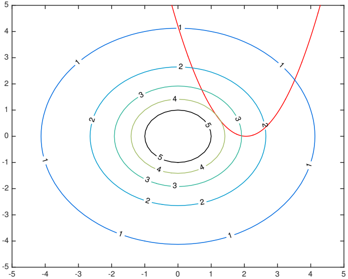
$$\text{maximize } f(x,y) \text{ subject to the constraint } g(x,y)=0.$$

Q: What are the graphs of the following equations?

$$z=f(x,y), \quad g(x,y)=0, \quad \text{and} \qquad k=f(x,y), \quad \text{where } k \text{ is a constant}$$

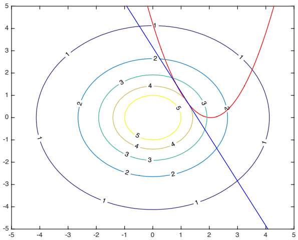
Q: Suppose $z=k$ is a local maximum, what is the relation between the graphs



- Notice at the point of the maximum, where the constraint curve and the level curve is met, the curves must share a common tangent line.

Q: What does it mean in terms of derivatives to have the same tangent line?



- Since $\nabla f$ is normal to the level curve

$$f(x, y) = k$$

and $\nabla g$ is normal to the constraint curve

$$g(x, y) = 0$$

- Thus the two gradients must be parallel,

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

for some scalar $\lambda$ at the point of tangency

$$(x_0, y_0)$$

- To prove this assertion, suppose a constrained local extremum exists and

the extremum occurs at $(x_0, y_0)$

- Further, the constraint equation $g(x, y) = 0$ can be smoothly parametrized

$$\mathbf{r}(s) = x(s)\mathbf{e}_x + y(s)\mathbf{e}_y$$

where $s$ is an arc length parameter with reference point $(x_0, y_0)$ at $s = 0$.

Q: Why $f(x(s), y(s)) = f(s)$ as a function of $s$ has a local extremum at $s = 0$?

$$\left.\frac{df}{ds}\right|_{s=0} = 0$$

- From the chain rule, the equation $\dfrac{df}{ds} = 0$ can be expressed as

$$\left.\frac{df}{ds}\right|_{s=0} = \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} \left.\frac{dx}{ds}\right|_{s=0} + \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} \left.\frac{dy}{ds}\right|_{s=0} = 0$$

- Consider the directional derivative of $f$ at $(x_0, y_0)$ in the direction given by

$$\mathbf{v} = \mathbf{r}'(s) = \mathbf{T}$$

- Hence the total derivative is simply the directional derivative at $(x_0, y_0)$,

$$f'_{\mathbf{T}} = \mathbf{T}(s=0) \cdot \nabla f(x_0, y_0) = \left.\frac{df}{ds}\right|_{s=0} = 0$$

- This implies that the gradient $\nabla f$ is either

zero or normal

to the constraint curve at a constrained local extremum $(x_0, y_0)$ since

$$\mathbf{T} \neq \mathbf{0}$$

Q: What happens if $\nabla f = \mathbf{0}$ at the extremum?

$$\mathbf{T}(s=0) \cdot \nabla f(x_0, y_0) = 0$$

- The constraint curve is a level curve for the function $g(x, y)$,

$$g(x, y) = 0$$

- Thus the gradient must orthogonal to the tangent at the point of tangency

$$\mathbf{T}(s = 0) \cdot \nabla g(x_0, y_0) = 0$$

hence we can formally conclude the vectors

$$\nabla g(x_0, y_0) \quad \text{and} \quad \nabla f(x_0, y_0) \quad \text{are parallel.}$$

- If $\nabla g(x_0, y_0) \neq \mathbf{0}$, then it follows that there exists some scalar $\lambda$ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

- This scalar $\lambda$ is called a Lagrange multiplier.

Q: There are two reasons for having the extra condition of

$$\nabla g(x_0, y_0) \neq \mathbf{0}$$

What is the obvious reason?

### The method of Lagrange Multipliers

Suppose that $f$ and $g$ are differentiable, and if

$$\nabla g \neq \mathbf{0} \qquad \text{when} \qquad g = 0$$

The local extremum values of $f$ subject to the constraint $g = 0$,

if there exist one,

can be found by finding the values of independent variables and $\lambda$ such that the following equations are simultaneously satisfied

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g = 0$$

### Exercise

*Use Lagrange multiplier to find the points in $\mathbb{R}^3$ on surface*

$$x^2 - y^2 - 1 = 0$$

*that are closest to the origin.*

## Exercise

(a) *Find the extreme values of the function*

$$f(x, y) = x^2 + 2y^2$$

*on the circle* $x^2 + y^2 = 1$.

- The method of Lagrange multipliers deals only equality constraints.

## Exercise

(b) *Find the extreme values of the function*

$$f(x, y) = x^2 + 2y^2$$

*on the disk* $x^2 + y^2 \leq 1$.

- Often we are required to find the extreme values of a differentiable function

$$f(x,y,z)$$

whose variables are subject to two or more constraints

$$g_1(x,y,z)=0 \qquad \text{and} \qquad g_2(x,y,z)=0$$

where $g_1$ and $g_2$ are differentiable, and $\nabla g_1$ not parallel to $\nabla g_2$, and nonzero

- Geometrically, the surfaces

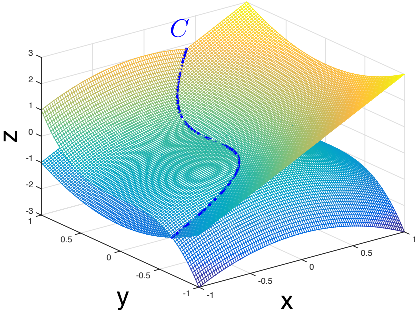
$$g_1 = 0 \quad \text{and} \quad g_2 = 0$$

intersect in a smooth curve $\mathcal{C}$

$$\mathbf{r}(t) = x(t)\mathbf{e}_x + y(t)\mathbf{e}_y + z(t)\mathbf{e}_z$$

- It is along this curve $\mathcal{C}$ we seek the points where $f$ has local extremum values relative to its other values on the curve $\mathcal{C}$.

**This is only the domain**



- The gradient $\nabla f$ at the extremum is normal to $\mathcal{C}$, that is,

$$\nabla f(x_0, y_0, z_0) \cdot \mathbf{T} = 0 \qquad \text{where} \quad \mathbf{T} = \frac{\mathbf{r}'}{|\mathbf{r}'|} \quad \text{at the extremum}$$

- Of course, the gradients $\nabla g_1(x_0, y_0, z_0)$ and $\nabla g_2(x_0, y_0, z_0)$ are normal to $\mathcal{C}$,

$$\begin{aligned} \nabla g_1(x_0, y_0, z_0) \cdot \mathbf{T} = 0 \\ \nabla g_2(x_0, y_0, z_0) \cdot \mathbf{T} = 0 \end{aligned} \implies \nabla f = \mu_1 \nabla g_1 + \mu_2 \nabla g_2 \quad \text{where } \mu_1, \mu_2 \in \mathbb{R}.$$

  since the cross product between the gradients is non-zero $(\nabla g_1) \times (\nabla g_2) \neq \mathbf{0}$

- So the constrained local maxima and minima of $f$ can be found by using

  two Lagrange multipliers $\mu_1$ and $\mu_2$.

### Exercise

(a) *The plane $x + y + z = 1$ cuts the cylinder $x^2 + y^2 = 1$ in an ellipse. Find the points on the ellipse that lie closest to and furthest from the origin.*

(b) *Apply the method of Lagrange multiplier to the post office problem earlier.*

## The method of Lagrange Multipliers

Suppose that $f$ and $g$ are differentiable, and if

$$\nabla g \neq \mathbf{0} \qquad \text{when} \qquad g = 0$$

The local extremum values of $f$ subject to the constraint $g = 0$,

if there exist one,

can be found by finding the values of independent variables and $\lambda$ such that the following equations are simultaneously satisfied

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g = 0$$

- Let us see the importance of the two conditions highlighted.
- Consider the following problem,

$$\begin{aligned} \text{max} \qquad & f(x,y) = x + y \\ \text{subject to} \qquad & g(x,y) = x^2 + y^2 = 1 \end{aligned}$$

- However, if we consider the following problem,

$$\begin{aligned} \text{max} \qquad & f(x,y) = x + y \\ \text{subject to} \qquad & g(x,y) = x^2 + y^2 = 0 \end{aligned}$$

- If we look closely, we notice the constrained set now is a single point, and

  the value $f(0,0) = 0$ is the maximum as well as the minimum.

- However, if we apply the method of Lagrange multiplier without checking

$$\nabla g \neq \mathbf{0}$$

- We have $\nabla f(0,0) = \lambda \nabla g(0,0)$ , so there is no value of $\lambda$, and no extremum

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- The moral is by neglecting the critical point of $g$ we could miss $f$'s extrema.

- Now consider the following problem,

$$\begin{aligned} \text{max} \qquad & f(x,y) = x + y \\ \text{subject to} \qquad & g(x,y) = xy - 16 = 0 \end{aligned}$$

- The method of Lagrange multiplier seems to suggest

$$\text{the value } f(4,4) = 8 \text{ is the maximum.}$$

  and

$$\text{the value } f(-4,-4) = -8 \text{ is the minimum.}$$

Q: However, is there any problem with those conclusions?

- The set is not bounded, so EVT is not applicable, thus LM is not applicable.
- In other words, $\nabla f = \lambda \nabla g$ is necessary for the occurrence of an extremum of $f$ subject to the conditions $\nabla g \neq \mathbf{0}$ and $g = 0$, but it is not sufficient.
- The lesson to be learnt is that if there is no solution to be found, the method of Lagrange multiplier might lead to incorrect conclusion.