MATH 2010E ADVANCED CALCULUS I
LECTURE 12

WING HONG TONY WONG

14.8 — Lagrange multipliers

In Section 14.7, we try to find extrema of f over a domain D in the following way.

(1) Find all critical points of f in $\text{int}(D)$.
(2) Classify the critical points of f using second derivative test if necessary.
(3) Find the extrema in $(\partial D) \cap D$.
(4) Compare the values found in (1) and (2) with those in (3) and decide the global extrema.

In this section, we will focus on the case where D is a level surface (or level curve). Since a level surface is expressed as $g(\mathbf{x}) = k$ for some constant k, we also call this type of problems "finding extrema with constraints".

Example 1. Find the minimal distance between the plane $x-2y+6z=4$ and the point $(3,-2,4)$.

Solution. We are minimizing the function $\phi(x,y,z) = \sqrt{(x-3)^2+(y+2)^2+(z-4)^2}$, which is equivalent to minimizing $f(x,y,z) = (x-3)^2+(y+2)^2+(z-4)^2$, subject to the constraint $x-2y+6z=4$.

We may rewrite the constraint as $x = 2y-6z+4$. Substituting back to $f(x,y,z)$, we are minimizing $\widetilde{f}(y,z) = (2y-6z+4-3)^2+(y+2)^2+(z-4)^2$.

$$\begin{cases} \widetilde{f}_y = 2(2y-6z+1)(2)+2(y+2) = 10y-24z+8 = 0 & (1) \\ \widetilde{f}_z = 2(2y-6z+1)(-6)+2(z-4) = -24y+74z-20 = 0 & (2) \end{cases}$$

$(1)\times 12 + (2)\times 5 \qquad -288z+96+370z-100 = 0 \ \Rightarrow\ 82z = 4 \ \Rightarrow\ z = \dfrac{2}{41}.$

Substitute back to (1), we get $10y - \dfrac{48}{41} + 8 = 0 \ \Rightarrow\ y = -\dfrac{28}{41}$.

$\widetilde{f}_{yy} = 10$, $\widetilde{f}_{yz} = -24$, and $\widetilde{f}_{zz} = 74$, so

$$D\widetilde{f} = \begin{pmatrix} 10 & -24 \\ -24 & 74 \end{pmatrix}.$$

Since $\widetilde{f}_{yy} = 10 > 0$ and $\det(D\widetilde{f}) = 10\times 74 - (-24)^2 = 164 > 0$, $(y,z) = \left(-\dfrac{28}{41}, \dfrac{2}{41}\right)$ is a local minimum of $\widetilde{f}(y,z)$. As $x = 2\left(-\dfrac{28}{41}\right) - 6\left(\dfrac{2}{41}\right) + 4 = \dfrac{96}{41}$ is well-defined, and the boundary of the plane $x-2y+6z=4$ is empty, the point $(x,y,z) = \dfrac{2}{41}(48,-14,1)$ is

Date: Thursday, 18th June, 2015.

the global minimum of f. The minimum distance is

$$f\left(\frac{2}{41}(48,-14,1)\right)=\sqrt{\left(-\frac{27}{41}\right)^2+\left(\frac{54}{41}\right)^2+\left(-\frac{162}{41}\right)^2}=\frac{9}{41}\sqrt{9+36+324}=\frac{27}{41}\sqrt{41}.$$

□

Example 2. Find the minimal distance between the surface $x^2-4yz-4z^2=-1$ and the origin.

Attempt. Similar to Example 1, we need to minimize $f(x,y,z)=x^2+y^2+z^2$, subject to the constraint $x^2-4yz-4z^2=-1$.

We may rewrite the constraint as $x^2=4yz+4z^2-1$. Substituting back to $f(x,y,z)$, we are minimizing $\widetilde{f}(y,z)=y^2+4yz+5z^2-1=(y+2z)^2+z^2-1$, which attains the minimum value -1 at $(y,z)=(0,0)$. However, $(y,z)=(0,0)$ implies that $x^2=-1$ which is impossible, so the global minimum value of $\widetilde{f}$ is not the global minimum value of f. Attempt failed...

□

Lemma 3. *Let $D\subseteq\mathbb{R}^n$ be a domain. Let $f:D\to\mathbb{R}$ be a differentiable real-valued function, and let $\mathbf{x}:I\to D$ be a smooth curve, where I is an interval of $\mathbb{R}$. If $\mathbf{a}\in\mathbf{x}(I)$ is a local extremum relative to the values of f on the curve $\mathbf{x}(I)$, then $\nabla f(\mathbf{a})$ is orthogonal to the tangent vector of the curve $\mathbf{x}(I)$ at $\mathbf{a}$.*

Solution. Let $t_0\in I$ be such that $\mathbf{x}(t_0)=\mathbf{a}$. Then $g(t)=f(\mathbf{x}(t))$, $t\in I$, attains a local extremum at t_0. By single-variable calculus, $g'(t_0)=0$. By chain rule,

$$0=g'(t_0)=\left.\frac{d}{dt}f(\mathbf{x}(t))\right|_{t=t_0}=Df(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}\cdot\mathbf{x}'(t_0)=\nabla f(\mathbf{a})\cdot\mathbf{x}'(t_0).$$

Therefore, $\nabla f(\mathbf{a})$ is orthogonal to the tangent vector of the curve $\mathbf{x}(I)$ at $\mathbf{a}$.

□

Let $\mathbf{a}$ be a local extremum of f subject to the level surface $D=\{\mathbf{x}\in\mathbb{R}^n: g(\mathbf{x})=k\}$. Then for all smooth curves $\mathbf{x}:I\to D$ such that $\mathbf{x}(t_0)=\mathbf{a}$, $\mathbf{a}\in\mathbf{x}(I)$ is a local extremum relative to the values of f on the curve $\mathbf{x}(I)$. By Lemma 3, $\nabla f(\mathbf{a})$ is orthogonal to all these curves at $\mathbf{a}$, and hence, orthogonal to the level surface $g(\mathbf{x})=k$ at $\mathbf{a}$. Recall that $\nabla g(\mathbf{a})$ is also orthogonal to the level surface $g(\mathbf{x})=k$ at $\mathbf{a}$ (if $\nabla g(\mathbf{a})\neq 0$). Therefore, $\nabla f(\mathbf{a})$ and $\nabla g(\mathbf{a})$ are parallel. This deduces the following result.

Theorem 4 (Lagrange multipliers)**.** *Let $\Omega\subseteq\mathbb{R}^n$ be a domain. Let $f,g:\Omega\to\mathbb{R}$ be two differentiable real-valued functions. Let $D=\{\mathbf{x}\in\Omega: g(\mathbf{x})=k\}$ be a level surface of g. If $\mathbf{a}\in D$ is a local extremum of f over D, and if $\nabla g(\mathbf{a})\neq\mathbf{0}$, then there exists $\lambda\in\mathbb{R}$ such that*

$$\nabla f(\mathbf{a})=\lambda\nabla g(\mathbf{a}).$$

(*Please note that λ can be* 0.)

The Lagrange multiplier method can also be understood in the following manner. If the level surface $f(\mathbf{x})=c$ and the level surface $g(\mathbf{x})=k$ cut across each other instead of tangent to each other, then if we make a small change from c to c', the level surface $f(\mathbf{x})=c'$ should still have nonempty intersection with the level surface $g(\mathbf{x})=k$. This means that c is not a local extremum. Hence, if we $f(\mathbf{x})$ attains a local extremum on the level surface $g(\mathbf{x})=k$ at $\mathbf{x}=\mathbf{a}$, then the level surfaces should be tangent to each other at $\mathbf{a}$. In other words, $\nabla f(\mathbf{a})$ should be parallel to $\nabla g(\mathbf{a})$.

Solution to Example 2. Let $\mathbf{a} = (a, b.c)$ be a local extremum of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint $g(x, y, z) = x^2 - 4yz - 4z^2 = -1$. By Lagrange multiplier method,

$$(2a, 2b, 2c) = \lambda(2a, -4c, -4b - 8c).$$

By $2a = 2\lambda a$, we have $a = 0$ or $\lambda = 1$. By $2b = -4\lambda c$, we have $b = -2\lambda c$. Substituting into $2c = -4\lambda b - 8\lambda c$, we have $c = 4\lambda(\lambda - 1)c$.

Case 1: $\lambda = 1$. Then $c = 0$, $b = -2$, and a is arbitrary. $g(a, -2, 0) = -1$ implies that $a^2 = -1$, impossible.

Case 2: $\lambda \neq 1$, and $c = 0$. Then $a = 0$, and $b = 0$. $g(0, 0, 0) = 0 \neq -1$, impossible.

Case 3: $\lambda \neq 1$, and $c \neq 0$. Then $a = 0$, and $4\lambda(\lambda - 1) = 1$, implying $\lambda = \dfrac{4 \pm \sqrt{32}}{8} = \dfrac{1 \pm \sqrt{2}}{2}$. This futher implies $b = (-1 \mp \sqrt{2})c$. As $g(a, b, c) = -1$, we have

$$\begin{aligned} 0^2 - 4(-1 \mp \sqrt{2})c^2 - 4c^2 &= -1 \\ \pm 4\sqrt{2}c^2 &= -1. \end{aligned}$$

So the positive sign is rejected, and $c = \pm\dfrac{1}{2\sqrt[4]{2}}$. Therefore,

$$(a, b, c) = \left(0, \pm\frac{-1 + \sqrt{2}}{2\sqrt[4]{2}}, \pm\frac{1}{2\sqrt[4]{2}}\right),$$

and the extreme value is

$$f(a, b, c) = \left(\pm\frac{-1 + \sqrt{2}}{2\sqrt[4]{2}}\right)^2 + \left(\pm\frac{1}{2\sqrt[4]{2}}\right)^2 = \frac{2 - \sqrt{2}}{2\sqrt{2}} = \frac{\sqrt{2} - 1}{2}.$$

To decide whether the local extreme value is a global maximum or minimum, we can try to understand the geometry of the level surface $g(x, y, z) = -1$.

The matrix for the quadratic portion of the equation is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & -2 & -4 \end{pmatrix}.$$

To find its eigenvalues,

$$\det \begin{pmatrix} 1 - \lambda & 0 & 0 \\ 0 & -\lambda & -2 \\ 0 & -2 & -4 - \lambda \end{pmatrix} = (1 - \lambda)(4\lambda + \lambda^2 - 4) = 0,$$

implying $\lambda = 1$ or $-2 \pm 2\sqrt{2}$. Therefore, there are two positive eigenvalues and one negative eigenvalue. Since $g(x, y, z) = -1 < 0$, this surface is a nondegenerated hyperboloid with two sheets.

Therefore, there is a global minimum for the distance between the surface and the origin, and there is no global maximum, and the desired minimum distance is $\dfrac{\sqrt{2} - 1}{2}$.

□

The geometric understanding at the end of the solution to Example 2 is crucial. This is because the solutions solved by Langrange multiplier method may not be a global extremum at all.

Example 5. Find the extreme values of $x + y$ subject to the constraint $xy = 1$.

Solution. Let $\mathbf{a} = (a, b)$ be a local extremum of $f(x, y) = x + y$ subject to the constraint $g(x, y) = xy = 1$. By Lagrange multiplier method,

$$(1, 1) = \lambda(b, a),$$

implying $a = b = \frac{1}{\lambda}$. So $g(a, b) = \frac{1}{\lambda^2} = 1$, i.e. $\lambda = \pm 1$. Therefore, $(a, b) = (1, 1)$ or $(a, b) = (-1, -1)$.

However, neither of them is a global extremum, since $g(x, y) = 1$ implies that $y = \frac{1}{x}$, so $x + y = x + \frac{1}{x}$ which can be arbitrary large (positive) and arbitrary small (negative). □

Example 6. Maximize and minimize $f(x, y) = xy$ subject to the constraint $x^2 + 2y^2 = 1$.

Solution. Let $\mathbf{a} = (a, b)$ be a local extremum of $f(x, y) = xy$ subject to the constraint $g(x, y) = x^2 + 2y^2 = 1$. By Lagrange multiplier method,

$$(b, a) = \lambda(2a, 4b).$$

Substituting $b = 2\lambda a$ into $a = 4\lambda b$, we have $a = 8\lambda^2 a$. If $a = 0$, then $b = 0$, and $g(a, b) = 0 \neq 1$, impossible. So $a \neq 0$ and $8\lambda^2 = 1$, i.e. $\lambda = \pm\frac{\sqrt{2}}{4}$. This means that $b = \pm\frac{\sqrt{2}}{2}a$.

$g\left(a, \pm\frac{\sqrt{2}}{2}a\right) = a^2 + 2\left(\pm\frac{\sqrt{2}}{2}a\right)^2 = 1$, which implies that $a = \pm\frac{\sqrt{2}}{2}$. So

(a, b)	$\left(\frac{\sqrt{2}}{2}, \frac{1}{2}\right)$	$\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$	$\left(-\frac{\sqrt{2}}{2}, \frac{1}{2}\right)$	$\left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$
$f(a, b)$	$\frac{\sqrt{2}}{4}$	$-\frac{\sqrt{2}}{4}$	$-\frac{\sqrt{2}}{4}$	$\frac{\sqrt{2}}{4}$

To decide whether the local extreme values are global maxima or global minima, note that the level surface $g(x, y) = x^2 + 2y^2 = 1$ is an ellipse, which is closed and bounded (i.e. compact). By the extreme value theorem (introduced in Lecture 11), there is a global maximum and global minimum of f on this level surface. Therefore, the desired maximum is $\frac{\sqrt{2}}{4}$, and the desired minimum is $-\frac{\sqrt{2}}{4}$. □

If $\mathbf{a}$ is a local extremum of f subject to multiple constraints $g_1(\mathbf{x}) = k_1$, $g_2(\mathbf{x}) = k_2$, $\ldots$, $g_j(\mathbf{x}) = k_j$, where $g_1, \ldots, g_j$ are differentiable, then

$$\nabla f(\mathbf{a}) = \lambda_1 \nabla g_1(\mathbf{a}) + \cdots + \lambda_j \nabla g_j(\mathbf{a})$$

if $\nabla g_1(\mathbf{a}), \ldots, \nabla g_j(\mathbf{a})$ are "linearly independent".

This can be understood in the following manner. Let the common intersection of all level surfaces $g_1(\mathbf{x}) = k_1$, $\ldots$, $g_j(\mathbf{x}) = k_j$ be Γ, which is of dimension $n - j$. Due to similar reasons mentioned in previous discussions, the level surface $f(\mathbf{x}) = c$ is tangent to Γ at $\mathbf{x} = \mathbf{a}$, where $c = f(\mathbf{a})$. In other words, $\nabla f(\mathbf{a})$ is orthogonal to Γ. Recall that Γ is of dimension $n - j$, so the orthogonal space of Γ is of dimension j, with $\nabla g_1(\mathbf{a}), \ldots, \nabla g_j(\mathbf{a})$ as its basis. Therefore, $\nabla f(\mathbf{a})$ is spanned by $\nabla g_1(\mathbf{a}), \ldots, \nabla g_j(\mathbf{a})$.

In a general Lagrange multiplier problem, we are solving the system of equations

$$\begin{cases} \nabla f(\mathbf{a}) &= \lambda_1 \nabla g_1(\mathbf{a}) + \cdots + \lambda_j \nabla g_j(\mathbf{a}) \\ g_1(\mathbf{x}) &= k_1 \\ &\vdots \\ g_j(\mathbf{x}) &= k_j \end{cases}.$$

The first equation yields n different equations, where n is the number of components. Together with the other j equations, there are $n+j$ equations. This matches perfectly with the number of variables, namely $a_1, \ldots, a_n, \lambda_1, \ldots, \lambda_j$.

Example 7. 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.

Solution. Let $\mathbf{a} = (a, b, c)$ be a local extremum of $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraints $g_1(x, y, z) = x + y + z = 1$ and $g_2(x, y, z) = x^2 + y^2 = 1$.

Since $\nabla g_1(\mathbf{a}) = (1, 1, 1)$ and $\nabla g_2(\mathbf{a}) = (2a, 2b, 0)$ are linearly independent when $(a, b, c) \neq (0, 0, 0)$, we can apply Lagrange multiplier method and have

$$(2a, 2b, 2c) = \lambda_1(1, 1, 1) + \lambda_2(2a, 2b, 0).$$

This implies $\lambda_1 = 2c$, $2a = 2c + 2\lambda_2 a$, $2b = 2c + 2\lambda_2 b$, i.e. $c = a(1 - \lambda_2) = b(1 - \lambda_2)$.
Case 1. If $\lambda_2 = 1$, then $c = 0$, $g_1(a, b, c) = a + b = 1$, and $g_2(a, b, c) = a^2 + b^2 = 1$. As a result,

$$ab = \frac{1}{2}[(a + b)^2 - (a^2 + b^2)] = 0,$$

implying that $(a, b, c) = (1, 0, 0)$ or $(a, b, c) = (0, 1, 0)$. In this case, $f(a, b, c) = a^2 + b^2 + c^2 = 1$.
Case 2. If $\lambda_2 \neq 1$, then $a = b$. $g_2(a, b, c) = a^2 + b^2 = 1$ implies that $a = b = \pm\frac{\sqrt{2}}{2}$. $g_1(a, b, c) = a + b + c = 1$ implies that $(a, b, c) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1 - \sqrt{2}\right)$ or $(a, b, c) = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 1 + \sqrt{2}\right)$. In this case, $f(a, b, c) = a^2 + b^2 + c^2 = 4 - 2\sqrt{2}$ or $4 + 2\sqrt{2}$ correspondingly.

Note that the intersection of the level surfaces $g_1(x, y, z) = 1$ and $g_2(x, y, z) = 1$ is an ellipse. Similar to Example 6, since an ellipse is closed and bounded (i.e. compact), there is a global maximum and global minimum of f on this ellipse.

Therefore, the points closest to the origin are $(1, 0, 0)$ and $(0, 1, 0)$ with minimum distance 1, and the point furthest from the origin is $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 1 + \sqrt{2}\right)$ with maximum distance $4 + 2\sqrt{2}$.

□

Example 8. Maximize $f(x, y, z) = x^2 + 3y - z^2$ subject to the constraints $g_1(x, y, z) = 2x - y = 0$ and $g_2(x, y, z) = y + z = 0$.

Solution. Let $\mathbf{a} = (a, b, c)$ be a local extremum of $f(x, y, z)$ subject to the given constraints. Since $\nabla g_1(\mathbf{a}) = (2, -1, 0)$ and $\nabla g_2(\mathbf{a}) = (0, 1, 1)$ are linearly independent, we can apply Lagrange multiplier method and have

$$(2a, 3, -2c) = \lambda_1(2, -1, 0) + \lambda_2(0, 1, 1).$$

Hence, $2\lambda_1 = 2a$, i.e. $\lambda_1 = a$, and $\lambda_2 = -2c$. Substituting into $-\lambda_1 + \lambda_2 = 3$, we have $-a - 2c = 3$.

$$\left(\begin{array}{ccc|c} 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & -2 & 3 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|c} -1 & 0 & -2 & 3 \\ 0 & 1 & 1 & 0 \\ 2 & -1 & 0 & 0 \end{array}\right) \xrightarrow{-R_1} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & 1 & 0 \\ 2 & -1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{R_3 - 2R_1} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -4 & 6 \end{array}\right) \xrightarrow{R_3 + R_2} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 6 \end{array}\right) \xrightarrow{-\frac{1}{3}R_3} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

$$\xrightarrow[R_2 - R_3]{R_1 - 2R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right).$$

Therefore, $(a, b, c) = (1, 2, -2)$, and $f(a, b, c) = 3$.

The intersection of the level surfaces $g_1(x, y, z) = 2x - y = 0$ and $g_2(x, y, z) = y + z = 0$ is a straight line. As there is only one local extremum, it has to be a global extremum as well. Since $(0, 0, 0)$ is in the intersection of the level surfaces, and $f(0, 0, 0) = 0 < 3$, $(a, b, c) = (1, 2, -2)$ is a global maximum. Therefore, the desired maximum is 3.

□

Example 9 (AM-GM inequality)**.** Show that if $x_1, x_2, \ldots, x_n \geq 0$, then

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \ldots, x_n}.$$

Solution. If $x_i = 0$ for some i, then the above inequality obviously hold.

Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a local extremum of $f(x_1, \ldots, x_n) = x_1 \ldots x_n$ subject to $g(x_1, \ldots, x_n) = \dfrac{x_1 + \cdots + x_n}{n} = r$, where $a_i \neq 0$ for all i. By Lagrange multiplier method,

$$a_1 \ldots a_n \left(\frac{1}{a_1}, \ldots, \frac{1}{a_n}\right) = \lambda \left(\frac{1}{n}, \ldots, \frac{1}{n}\right).$$

So $a_1 = \cdots = a_n$. Substituting into $g(a_1, \ldots, a_n) = \dfrac{a_1 + \cdots + a_n}{n} = r$, we have $(a_1, \ldots, a_n) = (r, \ldots, r)$, and $f(a_1, \ldots, a_n) = r^n$.

The level surface $g(x_1, \ldots, x_n) = \dfrac{x_1 + \cdots + x_n}{n} = r$ is an $(n-1)$-dimensional plane. As there is only one local extremum, it has to be a global extremum as well. Since $(nr, 0, \ldots, 0)$ is on the level surface, and $f(nr, 0, \ldots, 0) = 0 < r^n$, $(a_1, \ldots, a_n) = (r, \ldots, r)$ is a global maximum. Therefore, if $x_1, \ldots, x_n > 0$, let $\dfrac{x_1 + \cdots + x_n}{n} = r$. We have

$$x_1 \ldots x_n = f(x_1, \ldots, x_n) \leq f(r, \ldots, r) = r^n = \left(\frac{x_1 + \cdots + x_n}{n}\right)^n.$$

□