Lecture 16

Optimization with a Constraint

In this lecture we seek the max and min values of a function given a constraint on the independent variables of the function.

For example, what is the max and min values of \( f(x, y) = x + 3y - 4 \) on the circle \( x^2 + y^2 = 5 \)?

We will develop the ideas of the optimization technique, known as the **Method of Lagrange Multipliers**, by seeking the max and min values of \( f(x, y) \) given \( g(x, y) = c \). Note that the constraint equation \( g(x, y) = c \) is a level curve and serves as a restriction on the domain of \( f \).

We begin by parametrizing the level curve in vector form as \( \mathbf{r}(t) = (x(t), y(t)) \) so that the optimization problem is reduced to that of finding the max and min values of \( f(x(t), y(t)) \).

Note that \( f \) is now a function of a single variable \( t \) so that its critical points will occur when

\[
d/dt f(x(t), y(t)) = 0
\]
Applying the chain rule to the above derivative gives

\[
d/dt f(x(t), y(t)) = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} = 0
\]

\[
= \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle \cdot \langle \frac{dx}{dt}, \frac{dy}{dt} \rangle = 0
\]

Thus, \(\nabla f\) is orthogonal to \(r'(t)\) at a critical point.

Assuming both are non-zero.

We also know from previous lectures that \(r'(t)\) is tangent to the level curve and that \(\nabla g\) is orthogonal to the level curve.

We conclude that \(\nabla g\) and \(\nabla f\) are parallel at a critical point so that at these points we can write

\[
\nabla f = \lambda \nabla g
\]

where \(\lambda\) is a scalar known as a Lagrange multiplier.
Summarizing, we have found that the critical points for the optimization problem are the set of \((x, y)\) which simultaneously satisfy both

1. \(\nabla f(x, y) = \lambda \nabla g(x, y)\)

2. \(g(x, y) = c.\)

Having established the central ideas for the optimization technique we state for the record:

**Method of Lagrange Multipliers in Two Variables**

Let the function \(f\) and the constraining function \(g\) be differentiable on a region of \(R^2\) with \(\nabla g \neq 0\) on the curve \(g(x, y) = c.\) To locate the max and min values of \(f\) subject to the constraint \(g(x, y) = c,\) perform the following steps:

1. Find the values of \(x, y,\) and \(\lambda\) (if they exist) that satisfy the equations \(\nabla f = \lambda \nabla g\) and \(g(x, y) = c.\)

2. Evaluate \(f\) at the points \((x, y)\) found in Step 1; the largest and smallest values obtained are the max and min of \(f\) subject to the constraint.
Example: Find the max and min values of \( f(x, y) = x^2 - y^2 \) subject to \( x^2 + y^2 = 1 \).

\[
\nabla f = \langle 2x, -2y \rangle, \quad \nabla g = \langle 2x, 2y \rangle
\]

1. \( \nabla f = \lambda \nabla g \)
\[
2x = \lambda 2x, \\
-2y = \lambda 2y
\]
\[
\begin{align*}
x^2 + y^2 &= 1 \\
2x(1 - \lambda) &= 0 \\
2y(1 + \lambda) &= 0
\end{align*}
\]

Note: Simultaneous non-linear system

Consider \( 2x(1 - \lambda) = 0 \)

(a) If \( x = 0 \) then \( x^2 + y^2 = 1 \) \( \Rightarrow \) \( y = \pm 1 \) \n\[ \begin{align*}
(0, 1) \\
(0, -1)
\end{align*} \]

(b) If \( \lambda = 1 \), then \( 2y(1 + \lambda) = 0 \) \( \Rightarrow y = 0 \) \( \Rightarrow x = \pm 1 \) \n\[ \begin{align*}
(-1, 0) \\
(1, 0)
\end{align*} \]

Consider \( 2y(1 + \lambda) = 0 \)

(a) If \( y = 0 \) then \( x^2 + y^2 = 1 \) \( \Rightarrow x = \pm 1 \) \n\[ \begin{align*}
(-1, 0) \\
(1, 0)
\end{align*} \]

(b) If \( \lambda = -1 \) then \( 2x(1 - \lambda) = 0 \) \( \Rightarrow x = 0 \) \( \Rightarrow y = \pm 1 \) \n\[ \begin{align*}
(0, 1) \\
(0, -1)
\end{align*} \]

Thus 4 critical points are \( (0, 1), (0, -1), (1, 0), (-1, 0) \)

\[
\begin{align*}
f(0, 1) &= -1 \\
f(0, -1) &= -1 \\
f(1, 0) &= 1 \\
f(-1, 0) &= 1
\end{align*}
\]

max 1; min -1
Example: Find the max and min values of 
\[ f(x, y) = xy \] subject to 
\[ x^2 + y^2 - xy = 9. \]

\[ \nabla f = \langle y, x \rangle \quad \nabla g = \langle 2x-y, 2y-x \rangle \]

1. \[ \nabla f = x \nabla g \]

2. \[ x^2 + y^2 - xy = 9. \]

\[ y = x(2x-y) \quad \Rightarrow \quad x = x(2y-x) \]
\[ x^2 + y^2 - xy = 9. \]

Observe 1 & 2.
\[ y = x(2x-y) \quad \Rightarrow \quad xy = x(2x-y) \quad \Rightarrow \quad xy = x(2x-y) \]
\[ x = x(2y-x) \quad \Rightarrow \quad xy = x(2y-x) \]

Since \( \lambda \neq 0 \), \( x(2x-y) = y(2y-x) \) \( \Rightarrow \) \( y = \pm x. \)

Put this in equation 3

(a) \[ y = x \quad \Rightarrow \quad x^2 + x^2 - x^2 = 9 \quad \Rightarrow \quad x^2 = 9 \quad x = \pm 3 \]
\[ (3, 3); (-3, -3) \]

(b) \[ y = -x \quad \Rightarrow \quad 3x^2 = 9 \quad x = \pm \sqrt{3} \]
\[ (\sqrt{3}, -\sqrt{3}); (-\sqrt{3}, \sqrt{3}) \]

Now find \( f \) values.

\[ f(3, 3) = 9 \quad \text{max} \quad 9 \]
\[ f(-3, -3) = 9 \quad \text{min} \quad -3 \]
\[ f(\sqrt{3}, -\sqrt{3}) = -3 \]
\[ f(-\sqrt{3}, \sqrt{3}) = -3 \]
Example: Find the max and min values of 
\( f(x, y) = x^2 + y^2 \) subject to \( x^6 + y^6 = 1 \).

\[
\nabla f = \langle 2x, 2y \rangle \\
\nabla g = \langle 6x^5, 6y^5 \rangle
\]

\[
\begin{align*}
\text{1.} & \quad \nabla f = \lambda \nabla g \\
\text{2.} & \quad x^6 + y^6 = 1
\end{align*} \implies \begin{align*}
2x &= \lambda 6x^5 \\
2y &= \lambda 6y^5 \\
0 &= x(3\lambda x^4 - 1) \\
0 &= y(3\lambda y^4 - 1)
\end{align*}
\]

\[
\begin{align*}
x^6 + y^6 &= 1 \\
x^6 + y^6 &= 1
\end{align*}
\]

Consider the two equations

\[
\begin{align*}
0 &= x \left( 3\lambda x^4 - 1 \right) \\
0 &= y \left( 3\lambda y^4 - 1 \right)
\end{align*}
\]

(a) If \( x = 0 \) ; \( x^6 + y^6 = 1 \) \( \implies \) \( y = \pm 1 \) \( \implies \) \( (0, 1), (0, -1) \)

(b) If \( y = 0 \) ; \( x^6 + y^6 = 1 \) \( \implies \) \( x = \pm 1 \) \( \implies \) \( (1, 0), (-1, 0) \)

(c) If \( x \neq 0 \) \& \( y \neq 0 \) \( 3\lambda x^4 - 1 = 0 \) \( \implies \) \( 3\lambda y^4 - 1 \)

\[
\begin{align*}
x^4 &= y^4 \quad \implies \quad y = \pm x
\end{align*}
\]

If \( y = x \), \( x^6 + x^6 = 1 \) \( \implies \) \( x = \pm \frac{1}{\sqrt[6]{2}} \)

\[
\begin{align*}
\left( \frac{1}{\sqrt[6]{2}}, \frac{1}{\sqrt[6]{2}} \right), \left( -\frac{1}{\sqrt[6]{2}}, -\frac{1}{\sqrt[6]{2}} \right)
\end{align*}
\]

If \( y = -x \), \( x^6 + x^6 = 1 \) \( \implies \) \( x = \pm \frac{1}{\sqrt[6]{2}} \)

\[
\begin{align*}
\left( \frac{1}{\sqrt[6]{2}}, -\frac{1}{\sqrt[6]{2}} \right), \left( -\frac{1}{\sqrt[6]{2}}, \frac{1}{\sqrt[6]{2}} \right)
\end{align*}
\]

\[
\begin{align*}
f(0, 1) &= 1 \\
f\left( \frac{1}{\sqrt[6]{2}}, \frac{1}{\sqrt[6]{2}} \right) &= 2 \sqrt[3]{2} \\
f(0, -1) &= 1 \\
f\left( -\frac{1}{\sqrt[6]{2}}, -\frac{1}{\sqrt[6]{2}} \right) &= 2 \sqrt[3]{2} \\
f(-1, 0) &= \text{no} \\
f\left( \frac{1}{\sqrt[6]{2}}, -\frac{1}{\sqrt[6]{2}} \right) &= 2 \sqrt[3]{2} \\
f(1, 0) &= 1 \\
f\left( -\frac{1}{\sqrt[6]{2}}, \frac{1}{\sqrt[6]{2}} \right) &= 2 \sqrt[3]{2}
\end{align*}
\]

\[
\begin{align*}
\text{max} &= 2 \sqrt[3]{2} \\
\text{min} &= 1
\end{align*}
\]
Now consider the optimization of the function \( f(x, y, z) \) given the constraint \( g(x, y, z) = c \). In this case the constraint defines a level surface in \( R^3 \), and so the geometry of the problem has changed. The central ideas for the technique remain the same however. We can parametrize the level surface in vector form as \( \mathbf{r}(s, t) = \langle x(s, t), y(s, t), z(s, t) \rangle \) and seek the critical points of the function \( f(x(s, t), y(s, t), z(s, t)) \). At the critical points we find

\[
0 = \frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial s} = \langle f_x, f_y, f_z \rangle \cdot \langle x_s, y_s, z_s \rangle = \nabla f \cdot \frac{d\mathbf{r}}{ds}
\]

\[
0 = \frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial t} = \langle f_x, f_y, f_z \rangle \cdot \langle x_t, y_t, z_t \rangle = \nabla f \cdot \frac{d\mathbf{r}}{dt}
\]

Since the vectors \( \frac{\partial \mathbf{r}}{\partial s} \) and \( \frac{\partial \mathbf{r}}{\partial t} \) lie in the tangent plane of the level surface at a given point, \( \nabla f \) is orthogonal to the tangent plane of the surface at a critical point; we know from a previous lecture that \( \nabla g \) is also orthogonal to the tangent plane, thus at critical points \( \nabla f \) is parallel to \( \nabla g \), so that \( \nabla f = \lambda \nabla g \) at these points.

Thus the optimization procedure outlined for a function of two variables is unchanged if we increase to three independent variables. The computations become more difficult, however, since we introduce a fourth equation into the set which must be solved simultaneously.

\[
\begin{cases}
\nabla f = \lambda \nabla g \\
g(x, y, z) = c \\
\end{cases}
\]

\[
\begin{aligned}
f_x &= \lambda g_x \\
f_y &= \lambda g_y \\
f_z &= \lambda g_z \\
\end{aligned}
\]
Example: Find the max and min values of 
\[ f(x, y, z) = x + 3y - z \] subject to \[ x^2 + y^2 + z^2 = 4. \]

\[ \nabla f = \langle 1, 3, -1 \rangle \quad \nabla g = \langle 2x, 2y, 2z \rangle \]

1. \( \nabla f = \lambda \nabla g \)
   \[ \begin{align*}
   1 &= \lambda 2x \\
   3 &= \lambda 2y \\
   -1 &= \lambda 2z
   \end{align*} \]
   \( \lambda \neq 0 \)

\[ \frac{1}{2\lambda} = x \quad -6 \]
\[ \frac{3}{2\lambda} = y \quad -2 \]
\[ \frac{-1}{2\lambda} = z \quad -3 \]

2. \[ x^2 + y^2 + z^2 = 4 \]

Substituting 1, 2, 3, 4 in 4, we get \( \lambda = \pm \sqrt{11}/4. \)

\[ \lambda = \sqrt{11}/4 \Rightarrow \langle \frac{2}{\sqrt{11}}, \frac{6}{\sqrt{11}}, \frac{-2}{\sqrt{11}} \rangle \]

\[ \lambda = -\sqrt{11}/4 \Rightarrow \langle -\frac{2}{\sqrt{11}}, -\frac{6}{\sqrt{11}}, \frac{2}{\sqrt{11}} \rangle \]

\[ f(\langle \frac{2}{\sqrt{11}}, \frac{6}{\sqrt{11}}, \frac{-2}{\sqrt{11}} \rangle) = \frac{32}{\sqrt{11}} \quad \text{max} \]

\[ f(\langle -\frac{2}{\sqrt{11}}, -\frac{6}{\sqrt{11}}, \frac{2}{\sqrt{11}} \rangle) = -\frac{32}{\sqrt{11}} \quad \text{min} \]
Example. Find the min and max distance between the sphere \( x^2 + y^2 + z^2 = 20 \) and the point \((5, 6, 8)\).

\[
d = \sqrt{(x-5)^2 + (y-6)^2 + (z-8)^2}
\]

We can maximize \( f = D = d^2 = (x-5)^2 + (y-6)^2 + (z-8)^2 \)

\[
\nabla f = \langle 2(x-5), 2(y-6), 2(z-8) \rangle \quad \nabla g = \langle 2x, 2y, 2z \rangle
\]

\[
\nabla f = \lambda \nabla g \quad \left\{ \begin{array}{l}
2(x-5) = 2\lambda x \\
2(y-6) = 2\lambda y \\
2(z-8) = 2\lambda z \\
x^2 + y^2 + z^2 = 20
\end{array} \right.
\]

\[
\left\{ \begin{array}{l}
x-5 = \lambda x \\
y-6 = \lambda y \\
z-8 = \lambda z \\
x^2 + y^2 + z^2 = 20
\end{array} \right.
\]

\[
\sub (1, 2, 3) \text{ in (4)} \quad (\frac{5}{1-\lambda})^2 + (\frac{6}{1-\lambda})^2 + (\frac{8}{1-\lambda})^2 = 20
\]

\(
\lambda = -\frac{3}{2} \quad \Rightarrow \quad \langle 2, 12/5, 16/5 \rangle \quad \lambda = \frac{7}{2} \Rightarrow \quad \langle -2, -12/5, -16/5 \rangle
\)

\[
f(2, 12/5, 16/5) = 4.5 \\
f(-2, -12/5, -16/5) = 24.5
\]

\[
\min \ \text{dist} = 3\sqrt{5} \quad \max \ \text{dist} = 7\sqrt{5}
\]