ACTL3182 Lagrange Multipliers Revision

Andrew Wu

September 2020

This sheet provides a brief revision and some basic examples on Lagrange multipliers, a common method in constrained optimisation. In ACTL3182, the questions on Lagrange multipliers will resemble the lecture slides rather than these examples, which are only for your understanding.

1 Lagrange Multipliers

We begin simple with the one-constraint case that you learnt in MATH1251.

Definition 1.1: Gradient

The gradient (or derivative) of a function $f : \mathbb{R}^n \to \mathbb{R}$, is the vector of partial derivatives:

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right]^\top$$

- ∇f is equivalent to the vector derivative $\frac{\partial f}{\partial \boldsymbol{x}}$ when $\boldsymbol{x} \in \mathbb{R}^n$ i.e. when $\boldsymbol{x}$ is a vector of the same dimensions as the domain.
- ∇f is the direction in which f is increasing the most (steepest ascent). To see this, consider the Taylor-expansion
$$f(\boldsymbol{x} + \Delta\boldsymbol{x}) \approx f(\boldsymbol{x}) + \nabla f(\boldsymbol{x}) \cdot \Delta\boldsymbol{x},$$
where the dot product $\nabla f(\boldsymbol{x}) \cdot \Delta\boldsymbol{x} = \|\nabla f(\boldsymbol{x})\| \|\Delta\boldsymbol{x}\| \cos\theta$ is maximised when Δx is parallel to $\nabla f(\boldsymbol{x})$.

Theorem 1.1: Lagrange Multipliers (one constraint)

Suppose we want to optimise $f : \mathbb{R}^n \to \mathbb{R}$, subject to the constraint $g(\boldsymbol{x}) = C$, where $\boldsymbol{x} \in \mathbb{R}^n$ and $C \in \mathbb{R}$ is a constant. If $\nabla g \neq 0$, then f achieves its maxima/minima at the points where

$$\nabla f = \lambda \nabla g.$$

- Essentially, the gradients of f, g must be parallel for f to be maximised or minimised.
- The constant λ is called a **Lagrange Multiplier**

Note: $\nabla f = \lambda \nabla g$ is a **necessary but not a sufficient** condition for optima. That is, the equation can still hold at some point that is not actually a maximum or minimum. Therefore it is important to check the value of f at each point where $\nabla f = \lambda \nabla g$.

Example 1: One constraint

Find the points where the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$, is maximised on the ellipse

$$\frac{x^2}{2} + \frac{y^2}{3} = 1$$

Solution: Let g be the function given by

$$g(x, y) = \frac{x^2}{2} + \frac{y^2}{3}$$

Hence, we need to maximise f with respect to the constraint $g(x, y) = 1$. By the method of Lagrange multipliers, the maxima of f occur when

$$\begin{aligned}
&\nabla f = \lambda \nabla g, \quad \text{for some } \lambda \in \mathbb{R} \\
\iff &[y, x]^\top = \lambda \left[x, \frac{2y}{3}\right]^\top \\
\iff &y = \lambda x \quad (1), \quad \text{and} \quad x = \lambda \cdot \frac{2y}{3} \quad (2)
\end{aligned}$$

Dividing (1) by (2), we obtain

$$\begin{aligned}
&\frac{y}{x} = \frac{3x}{2y} \\
\iff &y^2 = \frac{3}{2}x^2 \qquad (3).
\end{aligned}$$

Substituting (3) into the constraint $g(x, y) = 1$,

$$\begin{aligned}
&\frac{x^2}{2} + \frac{\frac{3}{2}x^2}{3} = 1 \\
\iff &x = \pm 1
\end{aligned}$$

Substituting into (3), we obtain $y = \pm\sqrt{3/2}$. Clearly f is maximised when $xy > 0$, so f achieves its maximum at the points $(1, \sqrt{3/2})$ and $(-1, -\sqrt{3/2})$. Notice that we did not need to find the value of λ in this example, although sometimes it is easier to do so.

Now, we state the two-constraint theorem. Solving the two-constraint case by hand is complicated so the example has been moved to the appendix.

Theorem 1.2: Lagrange Multipliers (two constraints)

Let $f : \mathbb{R}^n \to \mathbb{R}$ denote the function to be optimised, subject to the constraints

$$\begin{aligned}
g_1(\boldsymbol{x}) = C_1 \\
g_2(\boldsymbol{x}) = C_2
\end{aligned}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ and C_1, C_2 are real constants. If $\nabla g_1, \nabla g_2 \neq 0$, then f achieves its maxima/minima at the points where

$$\nabla f = \lambda \nabla g_1 + \gamma \nabla g_2,$$

where $\lambda, \gamma \in \mathbb{R}$ are the Lagrange multipliers.

2 The Lagrangian

When we have more than one constraint, it is useful to repackage the constraints and the optimisation into a single function, which we can then optimise. This function is called the **Lagrangian**.

Theorem 2.1: Lagrange Multipliers (two constraints using the Lagrangian)

Let $f : \mathbb{R}^n \to \mathbb{R}$ denote the function to be optimised, subject to the constraints

$$\begin{aligned} g_1(\boldsymbol{x}) &= C_1 \\ g_2(\boldsymbol{x}) &= C_2 \end{aligned}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ and C_1, C_2 are real constants. The **Lagrangian** for the problem above is the function $\mathcal{L}$, where

$$\mathcal{L}(\boldsymbol{x}, \lambda, \gamma) = f(\boldsymbol{x}) + \lambda(C_1 - g_1(\boldsymbol{x})) + \gamma(C_2 - g_2(\boldsymbol{x}))$$

The function f achieves its maxima/minima when $\nabla\mathcal{L} = 0$, or equivalently,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}} = \mathbf{0}, \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 0, \qquad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \gamma} = 0.$$

where $\mathbf{0} \in \mathbb{R}^n$ is the zero vector.

The three conditions on the Lagrangian are equivalent to **Theorem 1.2**. The first condition equates the gradient of f to a linear combination of the gradients of g_1, g_2:

$$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}} = \mathbf{0} \\ \iff &\nabla f - \lambda \nabla g_1 - \gamma \nabla g_2 = \mathbf{0} \\ \iff &\nabla f = \lambda \nabla g_1 + \gamma \nabla g_2 \end{aligned}$$

The second and third conditions simply capture the constraints in the original problem:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda} &= C_1 - g_1(\boldsymbol{x}) = 0 \iff g_1(\boldsymbol{x}) = C_1 \\ \frac{\partial \mathcal{L}}{\partial \gamma} &= C_1 - g_2(\boldsymbol{x}) = 0 \iff g_2(\boldsymbol{x}) = C_1. \end{aligned}$$

Why use the Lagrangian? When solving problems by hand, it is not always useful. However, this form is signficantly better for implementing Lagrange multipliers using a computer.
This is because the computer can now treat the problem as one about finding where the gradient $\nabla\mathcal{L} = 0$, for which there exist nice optimisation techniques.
Conversely, if we were to use the form Theorem 1.2, the computer would need to solve a system of non-linear equations (eg. see example 2), which is a very difficult task for the computer.

ACTL3182 use: (Will make sense after the lecture) We want to minimise the portfolio variance $\sigma_P^2 = \boldsymbol{w}^\top \Sigma \boldsymbol{w}$ with respect to the constraints $\mathbf{1}^\top \boldsymbol{w} = 1$ and $\boldsymbol{z}^\top \boldsymbol{w} = \mu$. The Lagrangian is given by

$$\mathcal{L}(\boldsymbol{w}, \lambda, \gamma) = \frac{1}{2}\boldsymbol{w}^\top \Sigma \boldsymbol{w} + \lambda(1 - \mathbf{1}^\top \boldsymbol{w}) + \gamma(\mu - \boldsymbol{z}^\top \boldsymbol{w}),$$

and the required conditions are

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{w}} = \Sigma \boldsymbol{w} - \lambda \mathbf{1} - \gamma \boldsymbol{z} = \mathbf{0}, \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 1 - \mathbf{1}^\top \boldsymbol{w} = 0, \qquad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \gamma} = \mu - \boldsymbol{z}^\top \boldsymbol{w} = 0.$$

3 Appendix

Example 2: Two constraints

Find the point(s) where the function $f(x, y, z) = xyz$ attains its maximum value on the unit sphere $x^2 + y^2 + z^2 = 1$ and the plane $x - y + z = 1$.

Solution: Define the functions $g_1, g_2 : \mathbb{R}^3 \to \mathbb{R}$, where

$$\begin{aligned} g_1(x, y, z) &= x^2 + y^2 + z^2 \\ g_2(x, y, z) &= x - y + z. \end{aligned}$$

We want to maximise the function f with respect to the constraints $g_1(x, y, z) = 1$ and $g_2(x, y, z) = 1$. By the method of Lagrange multipliers,

$$\begin{aligned} &\nabla f = \lambda \nabla g_1 + \gamma \nabla g_2 \\ \iff &(yz, xz, xy)^\top = \lambda(2x, 2y, 2z)^\top + \gamma(1, -1, 1)^\top \end{aligned}$$

yielding the three equations:

$$yz = 2\lambda x + \gamma \tag{1}$$
$$xz = 2\lambda y - \gamma \tag{2}$$
$$xy = 2\lambda z + \gamma. \tag{3}$$

We eliminate γ by finding (1) + (2) and (1) + (3):

$$z(x + y) = 2\lambda(x + y) \iff (z - 2\lambda)(x + y) = 0 \tag{4}$$
$$x(y + z) = 2\lambda(y + z) \iff (x - 2\lambda)(y + z) = 0 \tag{5}$$

Case 1: If $x \neq -y$ and $y \neq -z$, then (4) and (5) reduce to

$$x = 2\lambda = z.$$

Substituting into $g_2(x, y, z) = 1$, we obtain $y = 2x - 1$. Substituting into $g_1(x, y, z) = 1$, we obtain

$$\begin{aligned} &2x^2 + (2x - 1)^2 = 1 \\ \iff &x(3x - 2) = 0. \end{aligned}$$

If $x = 0$, then $f = 0$ and if $x = 2/3$, then $y = 1/3, z = 2/3$ and $f = 4/27 > 0$. Thus, the maximum value of f given the constraints is $4/27$.

Case 2: If $x \neq -y$ but $y = -z$, then, $z = 2\lambda$ and $y = -2\lambda$. Substituting into $g_2(x, y, z) = 1$, we obtain $x = 1 - 4\lambda$. Substituting into $g_1(x, y, z) = 1$, we obtain $\lambda = 0, 1/3$. Ignoring $\lambda = 0$ (this gives another minimum) and taking $\lambda = 1/3$, we obtain $x = -1/3, y = -2/3, z = 2/3$.

Case 3: If $x = -y$ but $y \neq -z$, then $x = 2\lambda$, and similarly to case 2, we obtain $x = 2/3, y = -2/3, z = -1/3$.

Case 4: If $x = -y$ and $y = -z$, then substituting into $g_2(x, y, z) = 1$, we obtain $y = -1/3$, while substituting into $g_1(x, y, z) = 1$ yields $y = \pm 1/\sqrt{3}$. Hence, these conditions and the constraints are incompatible, so no maxima occur in this case.

Thus, f is maximised with repsect to the constraint $g_1(x, y, z) = 1$ and $g_2(x, y, z) = 1$ at the points

$$\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right), \qquad \left(-\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right), \qquad \text{and} \quad \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right).$$

Here are two exercises in case you want to practice applying Lagrange multipliers. Again, the questions in ACTL3182 are **not** like these exercises, but they may be helpful for your understanding. Worked solutions for these exercises will not be given.

Exercises

1. *Find the maximum and minimum values of $f(x, y) = x^2 + x + 2y^2$ on the unit circle $x^2 + y^2 = 1$.
2. **MATH1251 2012 Final 3iv):** Find the points on the ellipse $2x^2 - 4xy + 5y^2 = 54$ that are closest to the origin.

   Hint: Find two expressions for $\frac{y}{x}$ in terms of λ, then find the value(s) of λ.

Answers: 1. Minimum value is 0, maximum value is 9/4 2. $\left(\frac{-3}{\sqrt{5}}, \frac{6}{\sqrt{5}}\right)$ and $\left(\frac{3}{\sqrt{5}}, -\frac{6}{\sqrt{5}}\right)$

*Problem taken from James McKernan. *18.022 Calculus of Several Variables*. Fall 2010. Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu, License: Creative Commons BY-NC-SA.