10.018 Modelling Space and Systems

Lecture 4

Constrained Optimization: Lagrange Multipliers

Term 2, 2021

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Before we start....

To get the most out of this lecture, you should already be familiar with

1. Level curves and directional derivatives

as we will be going through

1. Constrained Optimization

Constrained Optimization

Suppose we have a **constrained** maximization problem:

$$\max f(\vec{x}), \quad \text{subject to} \quad \vec{x} \in R \subset \mathbb{R}^n.$$

R is some region in $\mathbb{R}^n$. One possible R is defined in such a way:

$$\max_{(x_1,\ldots,x_n)\in\mathbb{R}^n} f(x_1,\ldots,x_n), \quad \text{subject to} \quad \begin{aligned} g_1(x_1,\ldots,x_n) &= 0, \\ &\vdots \\ g_s(x_1,\ldots,x_n) &= 0, \\ h_1(x_1,\ldots,x_n) &\geq 0, \\ &\vdots \\ \text{and } h_t(x_1,\ldots,x_n) &\geq 0 \end{aligned}$$

Here $g_1(x_1,\ldots,x_n) = 0$, ..., $g_s(x_1,\ldots,x_n) = 0$ are the equality constraints and $h_1(x_1,\ldots,x_n) \geq 0$, ..., $h_t(x_1,\ldots,x_n) \geq 0$ are the inequality constraints that collectively define the feasible region R.

From constrained to unconstrained . . .

If we cannot solve the **constrained** problem directly, a natural approach is to study the **unconstrained** problem:

$$\max f(\vec{x}), \quad \vec{x} \in \mathbb{R}^n.$$

If a local maximum $\vec{x}^*$ for the unconstrained problem satisfies the constraints ($\vec{x}^* \in R$), it is also a local maximum for the original problem. This approach is called **relaxation** (see Lecture 3).

Recall: Optimization notation

From Lecture 3: Sometimes we will use this notation for optimization problems:

$$\max \; f(x,y)$$
$$\text{s.t.: } g(x,y) = 0.$$

Which means we want to find a maximum of $f(x,y)$ **subject to** constraint $g(x,y) = 0$.
OR

$$\min \; f(x,y)$$
$$\text{s.t.: } g(x,y) = 0.$$

Which means we want to find a minimum of $f(x,y)$ **subject to** constraint $g(x,y) = 0$.

. . . and back to constrained!

In some cases, we can try to solve the **constrained** problem directly:

$$\max f(\vec{x}), \quad \text{subject to} \quad \vec{x} \in R \subset \mathbb{R}^n.$$

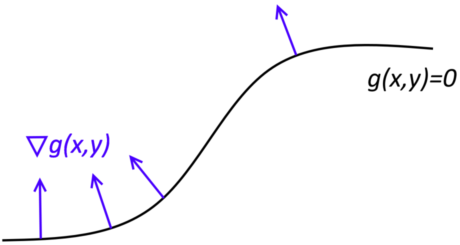
Today, we study a method to solve a **constrained** problem in the case of **equality constraints**. So the set R has this form:

$$R = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = 0, i = 1, \ldots, m.\}$$

Where $g_i(\vec{x}) = 0$ are **constraints**.

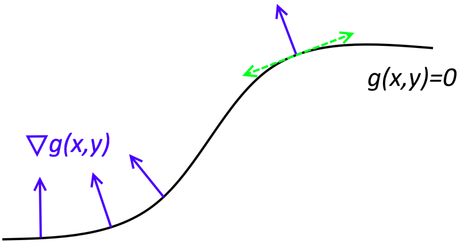
We can always make a constraint equal to zero on the right-hand side: for example, the constraint $\log xy = 5$ can be rewritten as $g(x, y) = \log xy - 5 = 0$.

Lagrange Multiplier



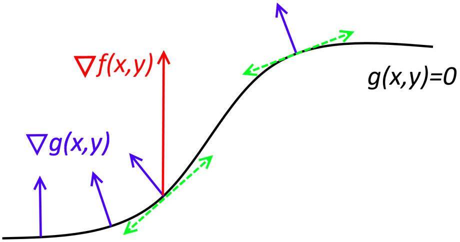
Consider the case of **one constraint**. The constraint $g(x, y) = 0$ defines a curve in the xy plane. The gradient $\nabla g(x, y)$ is perpendicular to the level curve $g(x, y) = 0$ at all points (see Week 1).

Lagrange Multiplier



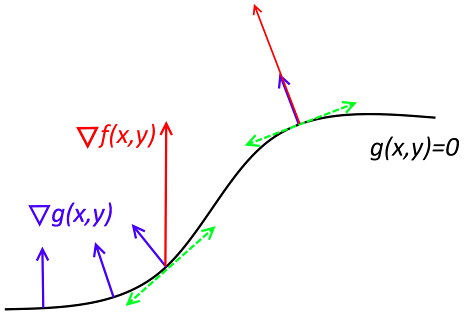
Because **we have to stay on the constraint**, the directions we can move from a point on the constraint are perpendicular to the gradient $\nabla g(x, y)$.

Lagrange Multiplier



We want to **maximize** function $f(x, y)$. The gradient $\nabla f(x, y)$ has a component along the direction we can move in. As this component is positive towards the right, we can increase the value of $f(x, y)$ while satisfying the constraint $g(x, y) = 0$ by moving to the right.

Lagrange Multiplier



The **only** place where we can't improve the value of $f(x, y)$ is where $\nabla f(x, y)$ is **parallel** to $\nabla g(x, y)$. This is the local optimum.

Equality constraints

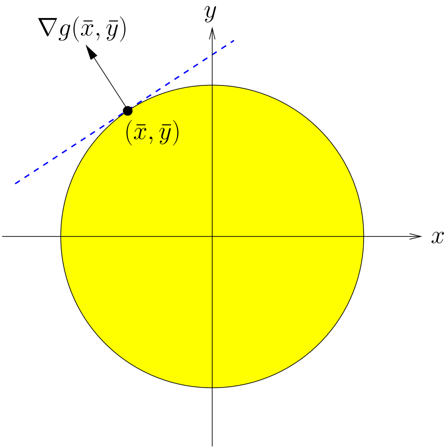
Let's look at the specific example:

$$\left.\begin{array}{rrcl} \max & x+y & & \\ \text{s.t.:} & x^2+y^2 & = & 1. \end{array}\right\} \qquad \left.\begin{array}{rrcl} \max & f(x,y) & & \\ \text{s.t.:} & g(x,y) & = & 0. \end{array}\right\}$$

We rewrite the **equality constraint** $x^2+y^2=1$ as $x^2+y^2-1=g(x,y)=0$ to give it a name.

Example

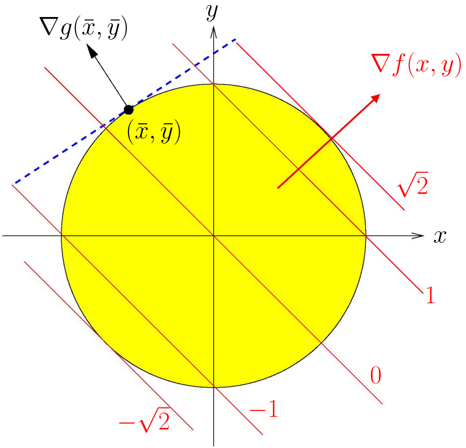
Maximize $f(x, y) = x + y$, subject to $g(x, y) = x^2 + y^2 - 1 = 0$



We want to increase the function value **while staying on the circle**.

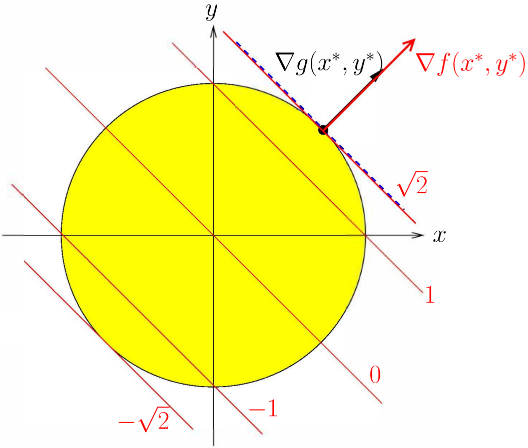
Optimality conditions: sketch

Maximize $f(x,y) = x + y$, subject to $g(x,y) = x^2 + y^2 - 1 = 0$



Optimality conditions: sketch

Maximize $f(x, y) = x + y$, subject to $g(x, y) = x^2 + y^2 - 1 = 0$



A point can only be optimal if $\nabla f // \nabla g$.

Solution

At the point $(x_1^*, x_2^*) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ we have:

$$\nabla g(x_1^*, x_2^*) = \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix}, \qquad \nabla f(x_1^*, x_2^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

We see graphically that this point is **optimal**.

This solves the initial problem:

$$\left.\begin{array}{lccc} \max & x + y & & \\ \text{s.t.:} & x^2 + y^2 & = & 1. \end{array}\right\}$$

Note that $(x_1^*, x_2^*) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and $\lambda^* = -\frac{1}{\sqrt{2}}$ satisfies: Lagrange multipliers for problems with one constraint.

Lagrange multipliers (One Constraint)

Theorem (Necessary condition for optimality)

Let $\vec{x}^*$ be a local maximum (or minimum) of f subject to $g(\vec{x}) = 0$. Then there exists a unique **Lagrange multiplier** $\lambda^* \in \mathbb{R}$ such that:

$$\nabla f(\vec{x}^*) + \lambda^* \nabla g(\vec{x}^*) = \vec{0}$$

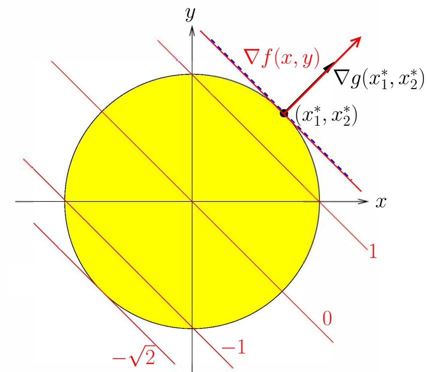
whenever $\nabla g(\vec{x}^*) \neq \vec{0}$ and $\vec{x}^*$ is not a boundary point of the constraint.

To solve a constraint optimization problem we need to find $\vec{x}^*, \lambda^*$ satisfying **both** equations:

$$\begin{aligned} \nabla f(\vec{x}) + \lambda \nabla g(\vec{x}) &= \vec{0} && \text{(Gradients are parallel)} \\ g(\vec{x}) &= 0 && \text{(Constraint is satisfied)} \end{aligned}$$

We can generalize these arguments for formulations with many equality constraints. (At the end of the lecture!)

Back to the problem



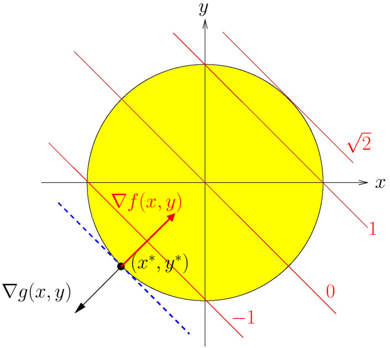
Let us understand the previous statement in our initial example. Recall that the optimum is $\vec{x}^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

$$\nabla g(x,y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

At the critical point

$$\nabla g\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} \quad \text{is parallel to} \quad \nabla f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Necessary is not sufficient!



Let us analyze another point, $\vec{x}^* = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

It also satisfies the Lagrange equations, but it is not a maximum. It is a minimum!

$$\nabla g(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = \begin{bmatrix} -\sqrt{2} \\ -\sqrt{2} \end{bmatrix} \quad \text{is parallel to} \quad \nabla f(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

In this class we do not cover sufficient conditions for min/max for Lagrange multiplier questions. In each individual problem you have to argue whether the point you found is a minimum or maximum.

Lagrange Multipliers (Several Constraints)

This discussion can be generalized. Consider a problem:

$$\left.\begin{array}{rrcl} \max & f(\vec{x}) & & \\ \forall i = 1, \ldots, m & g_i(\vec{x}) & = & 0 \\ & \vec{x} & \in & \mathbb{R}^n. \end{array}\right\}$$

Assume $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ are **continuously differentiable**. Then if $\vec{x}^*$ is a local maximum, there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that:

$$\nabla f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\vec{x}^*) = \vec{0}.$$

The scalars $\lambda_1, \ldots, \lambda_m$ are called **Lagrange multipliers**.

Necessary Conditions for Optimality

Lagrange multipliers for problems with several constraints.

Theorem

Let $\vec{x}^*$ be a local maximum (or minimum) of f subject to $g_i(\vec{x}) = 0, i = 1, \ldots, m$, and assume $\nabla g_1(\vec{x}^*), \ldots, \nabla g_m(\vec{x}^*)$ are **linearly independent**. Then there exist unique **Lagrange multipliers** $\lambda_1^*, \ldots, \lambda_m^* \in \mathbb{R}$ such that:

$$\nabla f(\vec{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\vec{x}^*) = \vec{0}$$

and $\vec{x}^*$ is not a boundary point of the constraints.

In this class we will see **at most two constraints**. In the case of two vectors **linear independence** means that the two vectors are not parallel, i.e. $\nabla g_1(\vec{x}^*) \neq k \nabla g_2(\vec{x}^*)$, for some scalar k.

Summary

- Analyzed **constrained** optimization problems with **equality** constraints.
- Presented the **Lagrange multiplier** theorem, and used it to find **candidate local minima/local maxima**.

Textbook: read Section 19.8, then try some of Exercises 19.8.1–19.8.23. You may discuss them on Piazza.