LECTURE 22: NON-LINEAR PROGRAMMING

Welcome to the world of non-linear programming (NLP), where everything that *can* go wrong *will* go wrong!

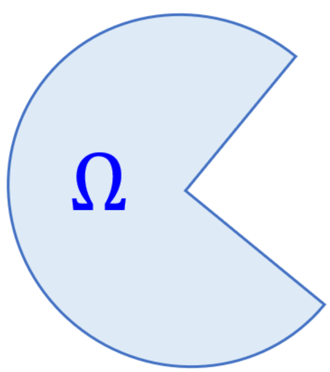
1. Introduction

A typical example would look as follows:

Example 1:

$$\begin{aligned} &\max z = (x_1)^3 + 4x_1x_2 + (x_2)^3 \\ \text{subject to } \ & (x_1)^2 + (x_2)^2 \leq 4 \\ & \ln(x_1) + \ln(x_2) \leq 3 \\ & x_1, x_2 \in \Omega \end{aligned}$$

Where Ω itself could be a nonlinear region, instead of $x_1 \geq 0$ and $x_2 \geq 0$



Date: Tuesday, November 29, 2022.

Here *everything* is nonlinear: The objective function, the constraints, and the domain Ω

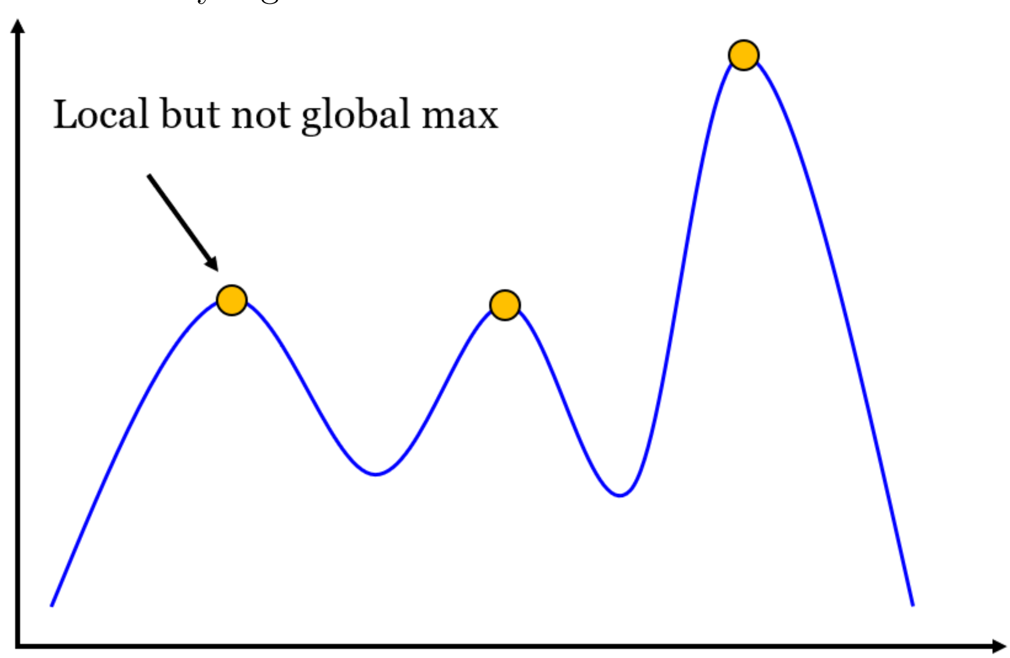
This is more general and more flexible... but much harder!

In fact, in the words of my advisor: "If you can solve all PDE, you can solve the universe!" The same thing goes with NLP. You can't have a unified theory of all NLP because that theory would have to hold for *all* models and *all* regions, including fractals!

In fact let's see some examples of what could go wrong

2. What can go wrong?

Problem 1: If the objective function is not convex, then a local max is not necessarily a global max

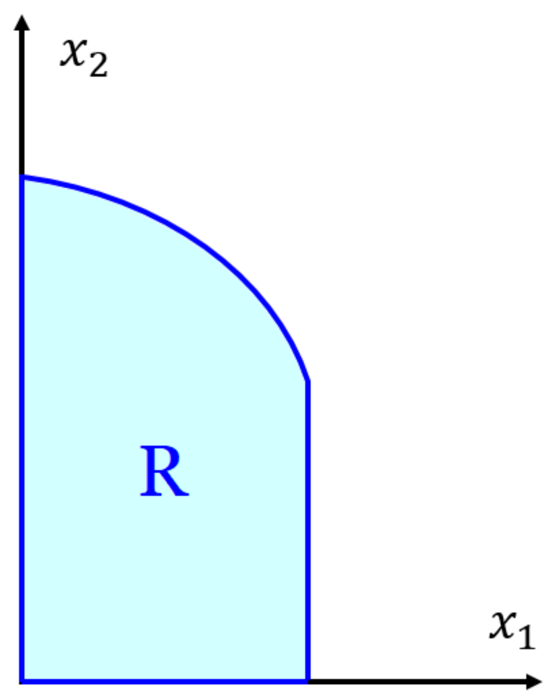


We have seen this phenomenon before, which was an issue even for linear programming (see below for definitions)

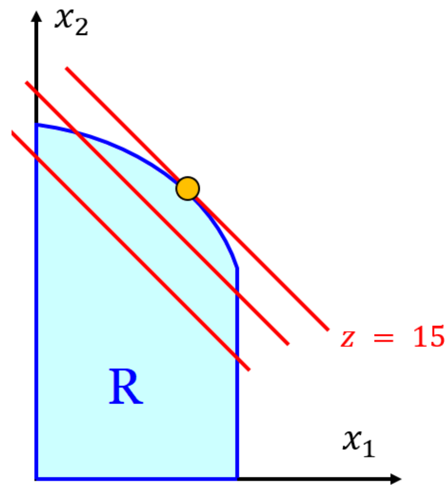
For LP, it was enough to check the vertices. This is not true for NLP:

Problem 2: If the feasible region is non-linear, then the max doesn't have to be at a vertex!

This is true even if the objective function is linear and the region is convex. In fact, suppose the feasible region R looks as follows:



Then, using the same method that motivated the simplex method, using level sets $z = 5$, $z = 10$, $z = 15$, we find that the max is not attained at a corner point:



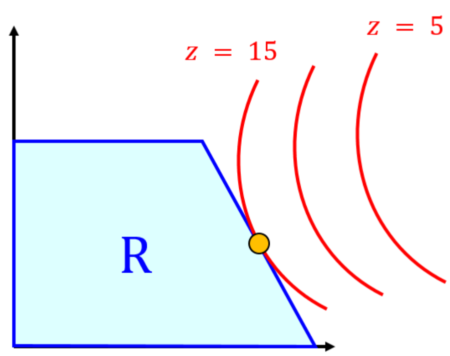
So our method of "checking the vertices" completely fails here!

What if the feasible region is linear but the objective function is non-linear?

Problem 3: Still false if R is linear!

Take for example the region below but z is some quadratic function

Then, even though R is a convex polygon, the max is once again not attained at a corner point



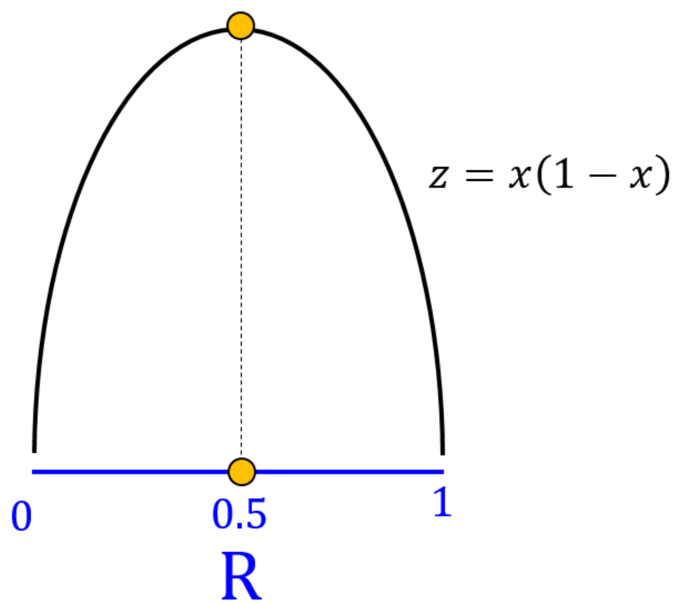
Problem 4: The max could be strictly inside R

Consider for instance the following 1D maximization problem:

Example 2:

$$\max z = x(1-x)$$
$$\text{subject to } 0 \leq x \leq 1$$

In this case the feasible region R is just $[0,1]$, whose boundary is 0 and 1, but the max is attained at $x = \frac{1}{2}$ which is strictly inside R



There are similar counterexamples as well in 2D, where R is a square and the max is attained strictly inside the square.

So even checking the whole boundary of R does not give us the max at all!

3. Applications

On the bright side, many problems can be naturally cast into NLP problems. They can be found throughout economics textbooks

Example 3:

You're the CEO of Peyamazon, and you're trying to figure out how much to sell books for

Suppose it costs \$20 to print each book, and the demand is $40-p$ where p is the price of the book.

What price should you sell the book for?

Decision Variables: $p =$ price per book that you charge

Objective Function: The revenue of each book is $p-20$ whereas the demand is $40-p$ so the total profit is

$$z = \text{ Revenue } \times \text{ Demand } = (p-20)(40-p)$$

Constraint: No constraints except the natural one $20 \leq p \leq 40$

NLP Problem:

$$\max z = (p-20)(40-p)$$
$$\text{subject to } 20 \leq p \leq 40$$

Note: This is an instance of an **unconstrained** maximization problem, where there are no (non-obvious) constraints

Note: You can replace $40-p$ by any demand function $D(p)$ as long as D is non-negative and decreasing

Example 4:

What is the key to happiness?

Chocolate and Vanilla ice cream, of course!

Suppose your utility/happiness function is given as follows

$$u(C,V) = C^{\frac{2}{3}} V^{\frac{1}{3}}$$

And moreover, chocolate ice cream costs \$2 a scoop, vanilla costs \$3 a scoop, and you only have \$20 available.

How many scoops of each do you pick to maximize your utility? Fractional scoops are allowed.

Note: u is called a **Cobb-Douglas** utility function. Even though each scoop of chocolate and vanilla makes you happier, the *additional* (marginal) happiness decreases with each scoop. So you get super happy with your first scoop of chocolate, but not as happy with your 10th scoop. Also $\frac{2}{3} > \frac{1}{3}$ means chocolate makes you happier than vanilla

Decision Variables: C, V = number of scoops of chocolate/vanilla

Objective Function:

$$u(C,V) = C^{\frac{2}{3}} V^{\frac{1}{3}}$$

Constraint: Since chocolate costs \$2 and vanilla costs \$3, and your budget is \$20, we get

$$2C + 3V \leq 20$$

NLP Problem:

$$\begin{aligned} \max\ u(C,V) &= C^{\frac{2}{3}}V^{\frac{1}{3}} \\ \text{subject to } 2C + 3V &\leq 20 \\ C, V &\geq 0 \end{aligned}$$

Note: Here the objective function is nonlinear, but the feasible region R is a convex polygon

4. Geometry of Feasible Regions

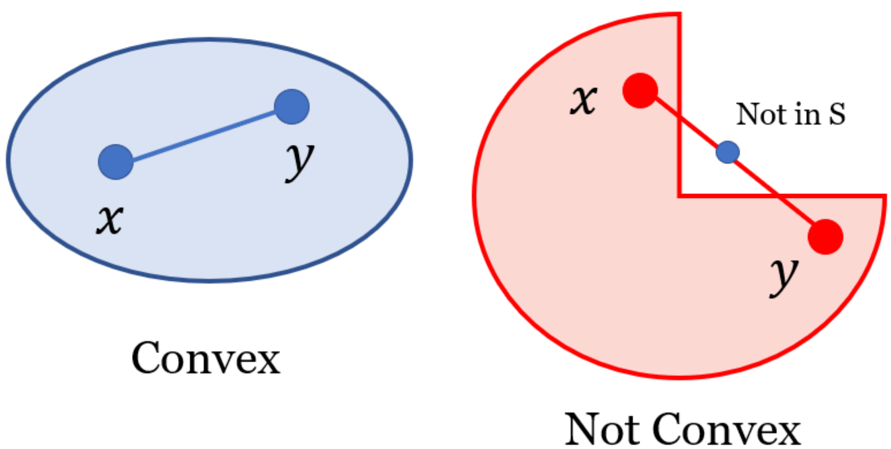
Once again, the key to solving NLP problems is... convexity!

Let's recall the definitions of convex sets and functions

Definition:

S is **convex** if for every $x, y \in S$, and every $0 \leq \lambda \leq 1$ we have

$$\lambda x + (1-\lambda)y \in S$$



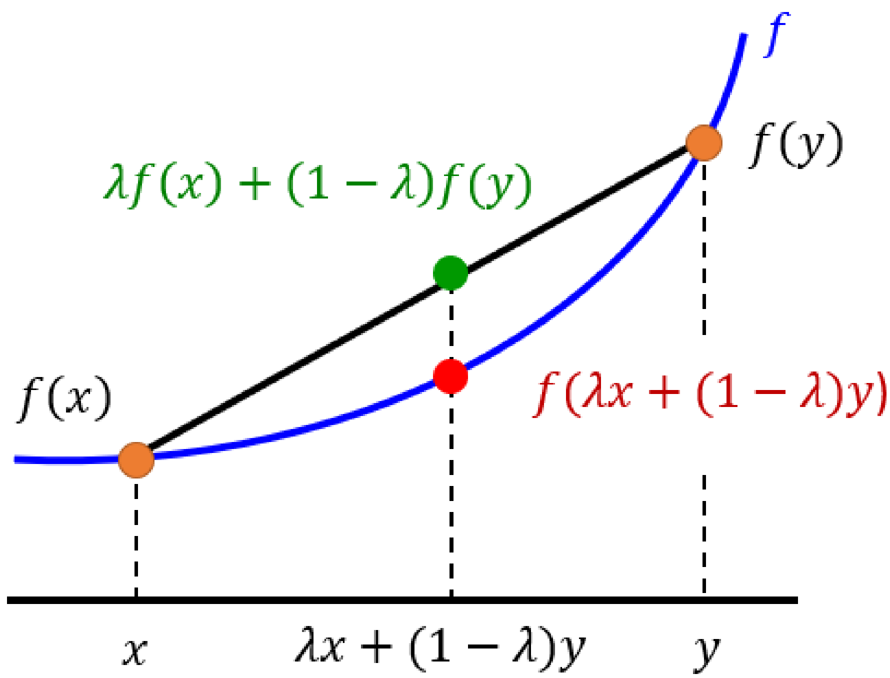
That is, for any points x, y in S, the line segment between x and y is also in S

Definition:

A function $f : S \to \mathbb{R}$ is convex if for every $x, y \in S$ and every $0 \leq \lambda \leq 1$, we have

$$f\left(\lambda x + (1-\lambda)y\right) \leq \lambda f(x) + (1-\lambda) f(y)$$

Interpretation: f lies below the segment connecting $f(x)$ and $f(y)$



1D Examples: $f(x) = x^2$ or $f(x) = e^x$ or $f(x) = -\ln(x)$ or $f(x) = |x|$

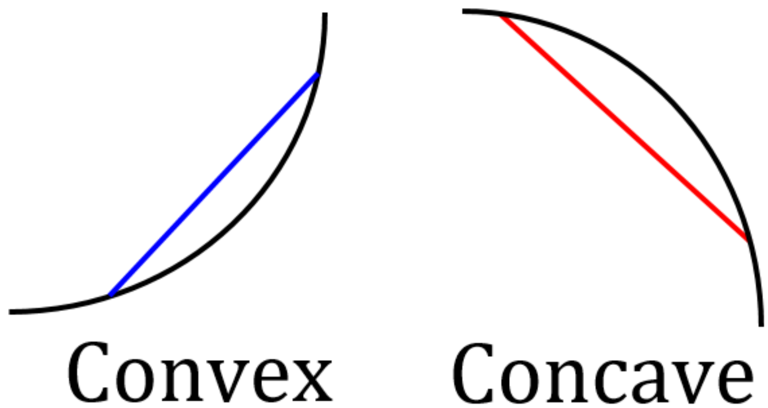
Definition:

f is **concave** if $-f$ is convex

Equivalently, we have

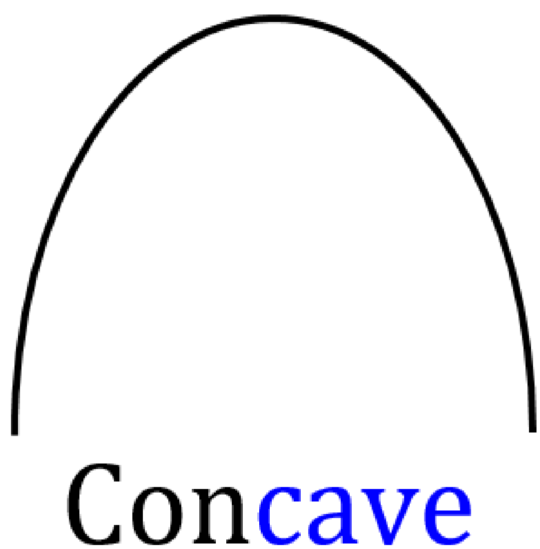
$$f\left(\lambda x+(1-\lambda) y\right) \geq \lambda f(x)+(1-\lambda) f(y)$$

That is, f lies **above** the segment connecting $f(x)$ and $f(y)$



Examples: $\ln(x)$ or $-x^2$

Mnemonic: Con*cave* like a *cave*

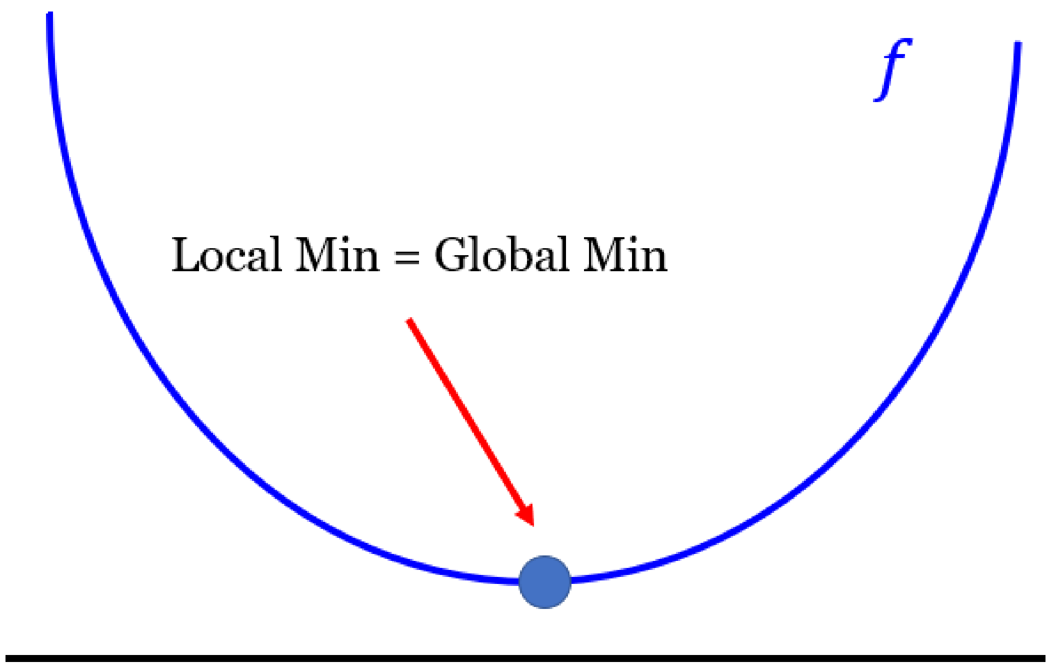


Note: Some functions that are neither convex and concave. But if a function is both convex and concave, then it is linear, that is $f(x) = a^T x + b$ for some vector a and some number b

Why is convexity useful?

Fact:

A *local* minimum of a convex function is a *global* minimum



Fact:

A local *maximum* of a *concave* function is a global *maximum*

(Follows because $\max f = -\min(-f)$ and $-f$ is convex)

Upshot: To solve our NLP problem (with min), if our feasible region and objective function is convex, it is enough to find a local min.

Hence, the problem boils down to:

(1) How to show f is convex? The definition above is not practical

(2) How to find local min?

5. The second-derivative test

In 1 dimensions, this question is much easier to solve, because of our powerful calculus tools.

How to show f is convex:

For this, we need the notion of concave up/down:

Definition (Concave up/down):

(1) f is concave up if $f''(x) \geq 0$ for all x

(2) f is concave down if $f''(x) \leq 0$ for all x

Fact:

IF f'' exists, then:

(1) Convex = Concave up

(2) Concave = Concave down

This gives us a *much* easier way of checking if f is convex.

Note: In general, if f'' doesn't exist, like $|x|$, you have to use the line segment definition above.

How to find local max/min:

For this, need the second-derivative test from calculus

Definition (Critical Point):

c is a **critical point** (CP) of f if $f'(c) = 0$ (or $f'(c)$ doesn't exist)

Second-Derivative Test:

Suppose c is a CP of f, then:

(1) If $f''(c) > 0$ then f has a local min at c

(2) If $f''(c) < 0$ then f has a local max at c

(And if $f''(c) = 0$ then the test is inconclusive)

To summarize:

Global Max/Min:

(1) If f is concave up and c is a CP then f has a global min at c

(2) If f is concave down and c is a CP then f has a global max at c

Next time: How to generalize this to higher dimensions?