Moving towards nonlinear optimization

CE 377K

April 9, 2015

ANNOUNCEMENTS

- HW 3 due Thursday
- Does everybody have a group?
- Project grading rubric and deadlines on the website

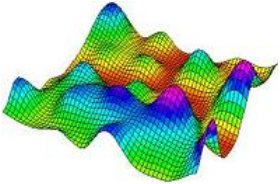
NONLINEAR OPTIMIZATION

The standard form for a nonlinear optimization problem is

$$
\begin{array}{lrl}
\min\limits_{\mathbf{x}} & f(\mathbf{x}) & \\
\text{s.t.} & g_1(\mathbf{x}) & \leq 0 \\
& \vdots & \\
& g_l(\mathbf{x}) & \leq 0 \\
& h_1(\mathbf{x}) & = 0 \\
& \vdots & \\
& h_m(\mathbf{x}) & = 0
\end{array}
$$

The objective function is to be minimized; all other constraints are of the form $\leq$ or $=$.

The *general* nonlinear optimization problem (where f, g, and h can be any functions whatever) is extremely difficult and probably impossible.



However, if the objective and constraints are "nice" functions, there are efficient algorithms for finding the global minimum.

At the start of this class, we saw some of these conditions (continuity, differentiability, unimodality, coercivity, boundedness, etc.)

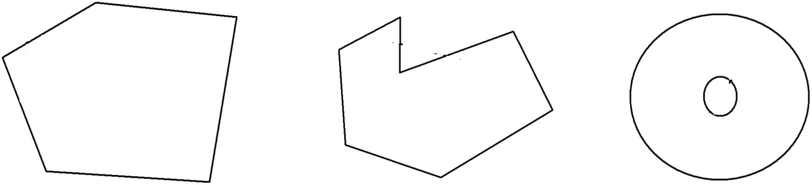
For nonlinear optimization problems the most important condition in practice is *convexity*.

There are actually *two* definitions of convexity, one applies to sets and the other applies to functions.

We will see that finding the global minimum of a convex function over a convex feasible set is achievable.

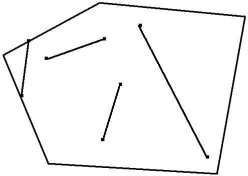
CONVEX SET

Intuitively, a convex set does not have any “holes” or “bites” in it.

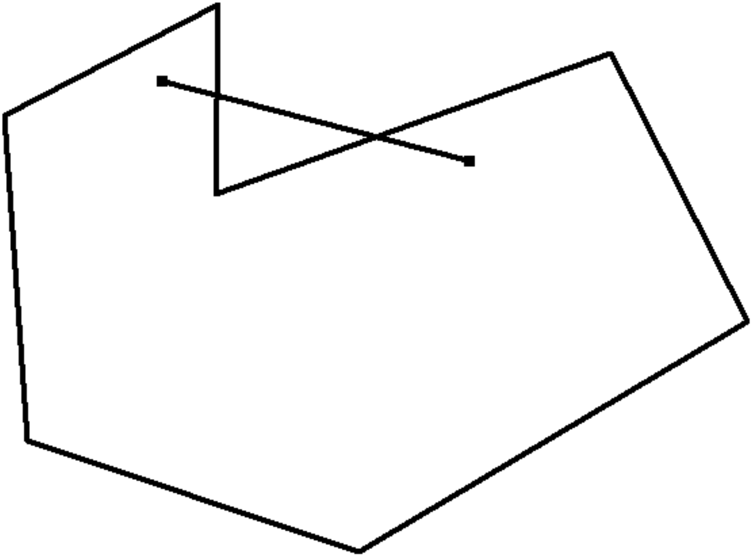


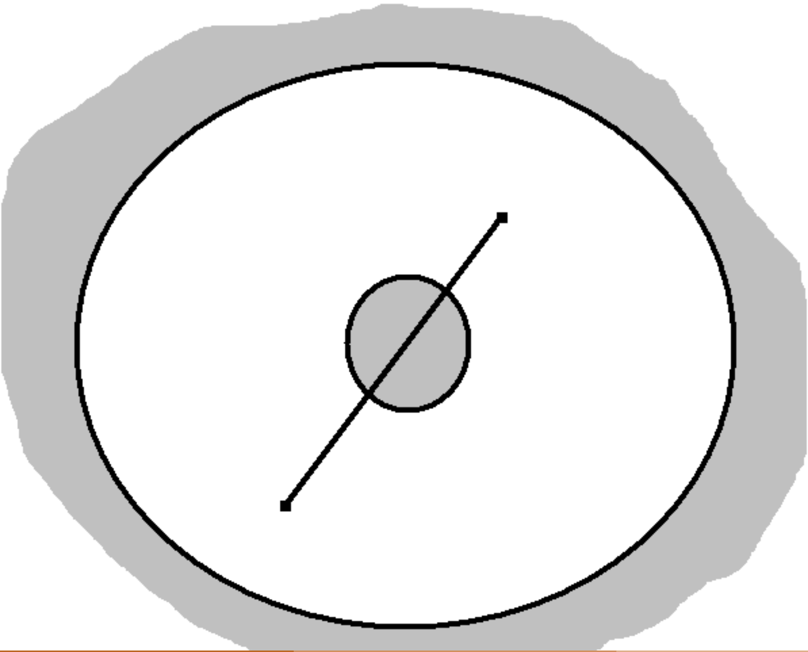
Convex Not convex Not convex

The more precise definition is that **for any two points in the set, the straight line connecting those two points also lies in the set.**



Specifically, the set X is convex if, for any $x_1 \in X$, $x_2 \in X$, and $\lambda \in [0, 1]$, the point $\lambda x_1 + (1 - \lambda)x_2 \in X$. (Such a point is a **convex combination** of x_1 and x_2.





Example

The one-dimensional set $X = \{x : x \geq 0\}$ is convex.

Pick any $x_1 \geq 0$, $x_2 \geq 0$, and $\lambda \in [0, 1]$.

Because all three of these are nonnegative, so is $\lambda x_1 + (1 - \lambda)x_2$.

Therefore the set is convex.

Example

The plane $X = \{(x, y, z) : 3x + 4y - 3z = 1\}$ is convex.

Pick any (x_1, y_1, z_1) and (x_2, y_2, z_2) in X, and any $\lambda \in [0, 1]$.

Then the convex combination is $(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2, \lambda z_1 + (1 - \lambda)z_2)$. Does this satisfy the conditions to be part of X?

We know $3x_1 + 4y_1 - 3z_1 = 1$ and $3x_2 + 4y_2 - 3z_2 = 1$.

Therefore $\lambda(3x_1 + 4y_1 - 3z_1) = \lambda$ and $(1-\lambda)(3x_2 + 4y_2 - 3z_2) = 1 - \lambda$.

Adding these shows that the convex combination $(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2, \lambda z_1 + (1-\lambda)z_2)$ also satisfies the equation of the plane, so it is convex.

Example

Is the region $X = \{(x, y) : x^2 + y^2 \geq 1\}$ convex?

The points $(1, 0)$ and $(-1, 0)$ are in X. Pick $\lambda = 1/2$.

The resulting point $(0, 0)$ is *not* in X, so X is not convex.

To show that a set is convex, you have to show that *every* convex combination of *every* two points in the set lie within the set. To show that a set is not convex, you only need one case where that is false.

CONVEX FUNCTIONS

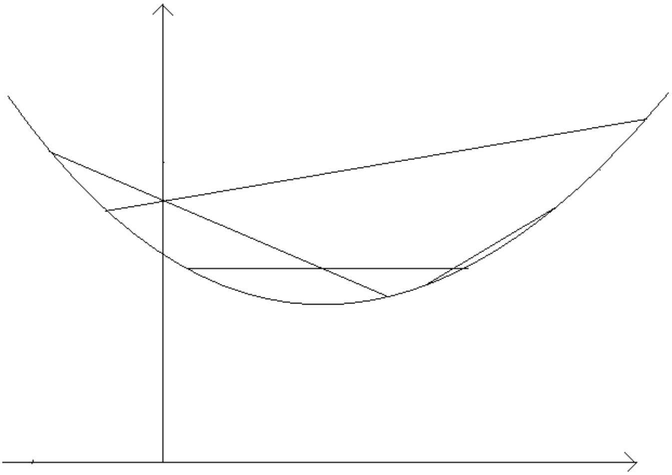
Function convexity is a bit different than set convexity.

We have already seen one definition of convexity early in the class (a one-dimensional, twice-differentiable function is convex if $f''(x) \geq 0$ everywhere.)

We will now generalize this definition to higher-dimension functions and to functions which are not twice differentiable.

Throughout this discussion, assume that the function's domain is a convex set.

Intuitively, a convex set lies below its secant lines.



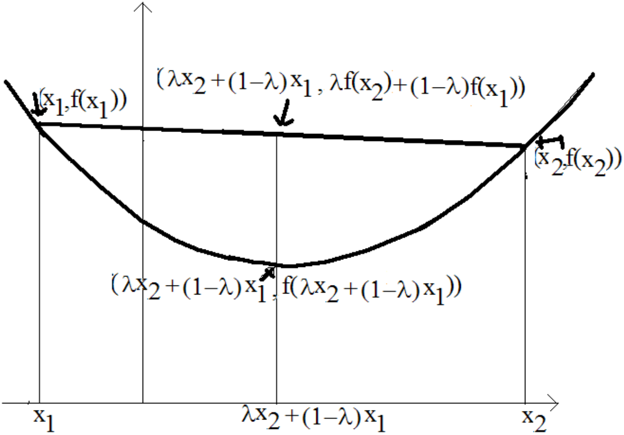
The mathematical way to express this is:

A function $f : X \to \mathbb{R}$ is *convex* if, for every $x_1, x_2 \in X$ and every $\lambda \in (0, 1)$,

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \tag{1}$$

Such a function is *strictly convex* if the $\leq$ can be replaced by $<$

Compare this definition with the figure:



Example

Is the function $f(x) = |x|$ convex? Is it strictly convex?

Pick any x_1, x_2, and $\lambda \in (0, 1)$.

$$f((1-\lambda)x_1 + \lambda x_2) = |(1-\lambda)x_1 + \lambda x_2|$$

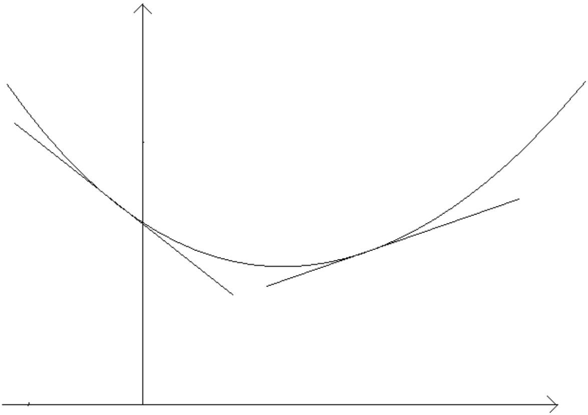
$\leq |(1-\lambda)x_1| + |\lambda x_2|$ by the triangle inequality

$= (1-\lambda)|x_1| + \lambda|x_2|$

$= (1-\lambda)f(x_1) + \lambda f(x_2)$.

This definition can be unwieldy to work with, so there are alternative characterizations.

If the function is differentiable, convexity can be characterized in terms of a function's *tangent* lines.



The function f is convex if it lies above all of its tangents.

Mathematically, if f is differentiable on its domain, then f is convex if and only if

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$$

for all $x_1, x_2 \in X$.

Example

Is x^2 convex?

Pick any x_1, x_2. Since $f'(x_1) = 2x_1$, we need to show that

$$x_2^2 \geq x_1^2 + 2x_1(x_2 - x_1)$$

This is equivalent to or

$$(x_1 - x_2)^2 \geq 0$$

which is always true, so f is convex.

If f is twice differentiable on its domain, then f is convex if and only if $f''(x) \geq 0$ everywhere.

Example: x^2 is convex because $f''(x) = 2 \geq 0$.

This is the definition we used earlier in the class.

When f is a function of multiple variables, the convexity conditions involving first and second derivatives must change.

The analogue of the first derivative is the *gradient* vector
$$\nabla f = \begin{bmatrix} \partial f/\partial x_1 & \partial f/\partial x_2 & \cdots & \partial f/\partial x_n \end{bmatrix}^T$$

The analogue of the second derivative is the *Hessian* matrix
$$Hf = \begin{bmatrix} \partial^2 f/\partial x_1^2 & \partial^2 f/\partial x_1 \partial x_2 & \cdots & \partial^2 f/\partial x_1 x_n \\ \partial^2 f/\partial x_2 \partial x_1 & \partial^2 f/\partial x_2^2 & \cdots & \partial^2 f/\partial x_2 x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 f/\partial x_n \partial x_1 & \partial^2 f/\partial x_n \partial x_2 & \cdots & \partial^2 f/\partial x_n^2 \end{bmatrix}$$

For twice-differentiable multidimensional functions, f is convex if any of these equivalent conditions are satisfied:

1. For all x_1 and x_2 in X,
$$f(\lambda x_2 + (1-\lambda)x_1) \leq \lambda f(x_2) + (1-\lambda)f(x_1)$$
2. For all x_1 and x_2 in X,
$$f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$$
3. For all x in X, $H(x)$ is positive semidefinite (that is, $y^T H(x) y \geq 0$ for all vectors y).

These conditions can be tedious to check. In this class I will not ask you to apply these definitions directly to multidimensional functions.

However, there are some facts which we can use (even in higher dimensions):

- Any linear function is convex.
- A nonnegative multiple of a convex function is convex.
- The sum of convex functions is convex.
- The composition of convex functions is convex.